

**INTER-CLASS SERVICE DIFFERENTIATION
AND INTRA-CLASS FAIRNESS
IN WDM OPTICAL BURST SWITCHING
NETWORKS**

TAN SIOK KHENG

(B.Eng. (Hons.), Sheffield University, UK)

**A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING
NATIONAL UNIVERSITY OF SINGAPORE**

January 2005

To Parents

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my mentor, Assistant Professor Mohan Gurusamy, for all the support, guidance and valuable discussion that made this work possible. Not only has he taught me the correct way of conducting research work, he has also inspired me on many levels as a researcher or a teacher. I am also grateful to Associate Professor Kee Chaing Chua for his valuable critiques and comments of my work.

I would also like to thank all the members of Open Source Software Lab (OSSL) who have made it such a great place to work. It has been a joyful moment working with them. I have also had a lot of support from the supervisor of OSSL, Associate Professor Bharadwaj Veeravalli and lab officer, Mr. David Koh. I would like to take this opportunity to express my appreciation to them.

I am especially grateful to my excellent parents and brothers for their endless love and encouragement. They have been a continual source of support and strength over many years.

The work in this thesis is supported in part the National University of Singapore Academic Research Grant No. R-263-000-173-112 and R-263-000-273-112.

Contents

Acknowledgements	ii
List of Figures	vii
List of Tables	xiii
Abstract	xv
1 Introduction	1
1.1 Overview of OBS	2
1.2 Motivation and Contribution	7
1.2.1 Fast and Efficient Burst Scheduling	7
1.2.2 Fairness in Multi-Hop WDM OBS Networks	8
1.2.3 Edge-to-Edge Proportional QoS	9
1.3 Organization of the Thesis	11
2 Background and Related Work	13

2.1	WDM Optical Networks	13
2.2	Transporting IP Traffic over WDM	14
2.3	Optical Switching Techniques	15
2.3.1	Optical Circuit Switching	16
2.3.2	Optical Packet Switching	16
2.3.3	Optical Burst Switching	16
2.4	OBS Networks	17
2.5	Optical Burst Switching Techniques	20
2.6	MPLS Framework for IP-over-WDM	23
2.7	Scheduling Algorithms	24
2.7.1	LAUC	26
2.7.2	LAUC-VF	27
2.7.3	PWA	28
2.7.4	BORA	28
2.8	QoS Provisioning	29
2.8.1	Offset-time based Service Differentiation	31
2.8.2	Segmentation based Service Differentiation	33
2.8.3	Scheduling based Service Differentiation	34
2.8.4	Preemption based Service Differentiation	34
2.8.5	Proportional Service Differentiation	35

2.8.6	Absolute Service Differentiation	37
2.9	Fairness	38
2.10	Summary	39
3	Burst Rescheduling Algorithms	40
3.1	Burst Rescheduling Techniques	42
3.1.1	Wavelength Reassignment	43
3.1.2	Last-hop FDL Reassignment	44
3.2	Burst Rescheduling Approaches	46
3.3	Burst Rescheduling Algorithms	46
3.3.1	On-Demand Burst Rescheduling (ODBR) Algorithm	48
3.3.2	Aggressive Burst Rescheduling (ABR) Algorithm	51
3.3.3	Burst Rescheduling with Wavelength and Last-hop FDL Reassignment (BR-WFR) Algorithm	55
3.4	Signalling Overhead	56
3.4.1	Signalling Overhead for ODBR	58
3.4.2	Signalling Overhead for ABR	59
3.4.3	Signalling Overhead for BR-WFR	59
3.5	Feasibility of Implementation	59
3.6	Performance Study	60
3.6.1	Simulation Model	60

3.6.2	Performance Metrics	61
3.6.3	Performance study of ODBR and ABR	62
3.6.4	Performance study of BR-WFR	63
3.6.5	Effect of Traffic Loading	68
3.6.6	Effect of FDL Buffer size	73
3.7	Summary	73
4	Offset Management for Fairness Improvement	76
4.1	Overview of LSOS	78
4.2	LSOS for Intra-class Fairness	81
4.2.1	Preliminaries	81
4.2.2	Computation of Link Scheduling Probabilities	85
4.2.3	Offset Selection	85
4.3	Performance Study	89
4.3.1	Performance of LSOS in a Classless Traffic Environment	90
4.3.2	Performance of LSOS in a Multi-class Environment	94
4.3.3	Effect of the Link-probing Phase Period on the Performance of LSOS	99
4.4	Summary	99
5	Edge-to-Edge Proportional QoS Provisioning	102
5.1	Supporting Proportional QoS with Extra Offset Times on a Single Link	103

5.1.1	Achievable Proportional Ratio Range - Two Classes	105
5.1.2	Achievable Proportional Ratio Range - Arbitrary Number of Classes . .	107
5.1.3	Achievable Proportional Ratio for a Given Offset Time	108
5.1.4	Numerical Results	110
5.2	Proposed FOTS Method	113
5.2.1	Overview of FOTS	114
5.2.2	Link State Collection	117
5.2.3	Traffic Measurement	118
5.2.4	Offset Time Selection	119
5.2.5	Supporting More than Two Traffic Classes	120
5.2.6	Convergence and Stability Issues	122
5.3	Performance Study	123
5.4	Summary	134
6	Conclusions	135
6.1	Research Contribution	136
6.2	Future Work	138
	Bibliography	140
	Author's Publications	150

List of Figures

1.1	Various quality of service issues in WDM OBS networks	11
2.1	Possible protocol stack options for IP-over-WDM	15
2.2	Separation of control channel(s) and data channel(s) in OBS.	18
2.3	An optical burst switching network.	19
2.4	General architecture of an OBS node	19
2.5	The use of offset time and immediate reservation in JIT.	23
2.6	The use of offset time and delayed reservation in JET.	23
2.7	Illustration of LAUC and LAUC-VF.	27
3.1	Illustration of the benefit of burst rescheduling. (a) Both LAUC and LAUC-VF fail to schedule the new burst. (b) The new burst is scheduled by rescheduling burst 3.	43
3.2	Illustration of the benefit of wavelength reassignment. (a) LAUC fails to schedule burst 7. (b) Burst 7 can be scheduled by using wavelength reassignment.	44

3.3	Illustration of the benefit of burst rescheduling with FDL reassignment. (a) LAUC fails to schedule the new burst, wavelength reassignment does not help. (b) The new burst is scheduled by allowing FDL reassignment.	45
3.4	Illustration of multi-level rescheduling. (a) No wavelength is available for new burst. (b) Rescheduling of burst 4 from W_2 to W_3 followed by rescheduling of burst 2 from W_1 to W_2 frees W_1 to accommodate new burst.	47
3.5	Illustration of ODBR. (a) A situation wherein the new burst can not be scheduled. (b) The last burst on W_3 is moved to W_2 to accommodate the new burst on W_3	50
3.6	Illustration of a situation wherein LAUC, ODBR and LAUC-VF fail to schedule new burst 6.	53
3.7	Illustration of working of ABR. (a) New burst 4 is assigned to W_2 . (b) Last burst from W_1 is rescheduled to W_2 . (c) Burst 5 is assigned to W_2 . (d) Burst 6 will be able to be scheduled to W_1	54
3.8	Performance of overall traffic for various algorithms under different traffic loading.	64
3.9	Performance of class 1 traffic for various algorithms under different traffic loading.	64
3.10	Performance of class 2 traffic for various algorithms under different traffic loading.	65
3.11	Performance improvement of overall traffic for various algorithms under different traffic loading.	65
3.12	Performance improvement of class 1 traffic for various algorithms under different traffic loading.	66
3.13	Performance improvement of class 2 traffic for various algorithms under different traffic loading.	66

3.14	Effectiveness of overall traffic for ODBR and ABR under different traffic loading.	67
3.15	Effectiveness of class 1 traffic for ODBR and ABR under different traffic loading.	67
3.16	Effectiveness of class 2 traffic for ODBR and ABR under different traffic loading.	68
3.17	Performance of overall (class 1 and class 2) bursts for varying traffic load. . . .	70
3.18	Performance of class 1 bursts for varying traffic load.	70
3.19	Performance of class 2 bursts with varying traffic load.	71
3.20	Performance improvement achieved by BR-WFR, BR-WR, and LAUC-VF over LAUC for overall bursts for varying traffic load.	71
3.21	Performance improvement achieved by BR-WFR, BR-WR, and LAUC-VF over LAUC for class 1 bursts with varying traffic load.	72
3.22	Performance improvement achieved by BR-WFR, BR-WR, and LAUC-VF over LAUC for class 2 bursts with varying traffic load.	72
3.23	Performance of class 1 bursts for varying FDL size.	73
3.24	Performance of class 2 bursts for varying FDL size.	74
4.1	Link states on a 2-hop path.	80
4.2	Division of offset time into frames for different priority classes of traffic. . . .	82
4.3	Illustration of link state tables generated at nodes.	87
4.4	14-node NSFNET.	91
4.5	Dropping performance vs. hop length for classless traffic with identical traffic demand.	92

4.6	Dropping performance vs. hop length for classless traffic with non-identical traffic demand.	93
4.7	Dropping performance vs. hop length for class 1 traffic with identical traffic demand.	95
4.8	Dropping performance vs. hop length for class 2 traffic with identical traffic demand.	96
4.9	Dropping performance vs. hop length for class 1 traffic with non-identical traffic demand.	96
4.10	Dropping performance vs. hop length for class 2 traffic with non-identical traffic demand.	97
4.11	Standard deviation vs. link probing period for classless traffic.	100
4.12	Standard deviation vs. link probing period for class 1 traffic.	100
4.13	Standard deviation vs. link probing period for class 2 traffic.	101
5.1	Upper bound of the achievable proportional ratio, $R_{1,2}^U$ (with complete isolation) for traffic composition 50H-50L.	111
5.2	Upper bound of the achievable proportional ratio, $R_{1,2}^U$ (with complete isolation) for traffic composition 30H-70L.	112
5.3	Achievable proportional ratio, $R_{1,2}''$ (without complete isolation) for traffic composition 50H-50L.	113
5.4	Achievable proportional ratio, $R_{1,2}''$ (without complete isolation) for traffic composition 30H-70L.	114

5.5	Illustration of FOTS with probing for traffic measurement collection and traffic measurement period on time axis.	116
5.6	Probe packet format for link state collection	116
5.7	Proportional ratio achieved between class 1 and class 2, with $R_{1,2}^d = 3$, $T_p = 50$ msec and arrival rate of 0.1 <i>bursts</i> / μ sec.	124
5.8	Proportional ratio achieved between class 1 and class 2, with $R_{1,2}^d = 3$, $T_p = 100$ msec and arrival rate of 0.1 <i>bursts</i> / μ sec.	125
5.9	Proportional ratio achieved between class 2 and class 3, with $R_{1,2}^d = 3$, $T_p = 50$ msec and arrival rate of 0.1 <i>bursts</i> / μ sec.	125
5.10	Proportional ratio achieved between class 2 and class 3, with $R_{1,2}^d = 3$, $T_p = 100$ msec and arrival rate of 0.1 <i>bursts</i> / μ sec.	126
5.11	Proportional ratio achieved between class 1 and class 2, with $R_{1,2}^d = 3$, $T_p = 50$ msec and arrival rate of 0.2 <i>bursts</i> / μ sec.	127
5.12	Proportional ratio achieved between class 1 and class 2, with $R_{1,2}^d = 3$, $T_p = 100$ msec and arrival rate of 0.2 <i>bursts</i> / μ sec.	127
5.13	Proportional ratio achieved between class 2 and class 3, with $R_{2,3}^d = 5$, $T_p = 50$ msec and arrival rate of 0.2 <i>bursts</i> / μ sec.	128
5.14	Proportional ratio achieved between class 2 and class 3, with $R_{2,3}^d = 5$, $T_p = 100$ msec and arrival rate of 0.2 <i>bursts</i> / μ sec.	128
5.15	Proportional ratio achieved between class 1 and class 2, with $R_{1,2}^d = 3$, and arrival rate of 0.1 <i>bursts</i> / μ sec.	129

5.16	Proportional ratio achieved between class 1 and class 2, with $R_{1,2}^d = 3$, and arrival rate of 0.2 <i>bursts/μsec.</i>	130
5.17	Proportional ratio achieved between class 2 and class 3, with $R_{2,3}^d = 5$, and arrival rate of 0.1 <i>bursts/μsec.</i>	130
5.18	Proportional ratio achieved between class 2 and class 3, with $R_{2,3}^d = 5$, and arrival rate of 0.2 <i>bursts/μsec.</i>	131
5.19	Average offset time needed for class 2 traffic with $R_{1,2}^d = 3$, and arrival rate of 0.1 <i>bursts/μsec.</i>	132
5.20	Average offset time needed for class 2 traffic with $R_{1,2}^d = 3$, and arrival rate of 0.2 <i>bursts/μsec.</i>	132
5.21	Average offset time needed for class 3 traffic with $R_{2,3}^d = 5$, and arrival rate of 0.1 <i>bursts/μsec.</i>	133
5.22	Average offset time needed for class 3 traffic with $R_{2,3}^d = 5$, and arrival rate of 0.2 <i>bursts/μsec.</i>	133

List of Tables

3.1	ODBR algorithm	49
3.2	ABR algorithm	52
3.3	BR-WFR algorithm	57
4.1	2-hop path scheduling probability for different offset time values.	81
4.2	Offset Time Assignment to Different Priority Classes	84
4.3	Computation of Link Scheduling Probabilities	86
4.4	Offset Time Selection	88
4.5	A-LSOS and 1-LSOS Path Scheduling Probability of path 1-2-3.	89
4.6	Standard deviation in burst dropping probabilities with different hop lengths for classless environment.	92
4.7	Mean offset time (in μs) needed for A-LSOS, 1-LSOS, JET, and JET-FA for classless environment.	93
4.8	Standard deviation in burst dropping probabilities of traffic with different hop lengths.	97

4.9	The mean offset time (in μs) needed for A-LSOS, 1-LSOS, and pJET in multi-class traffic with different hop lengths.	98
5.1	Offset time table carried by a probe packet.	118
5.2	Offset Time Selection in LSOS	121

Abstract

Wavelength division multiplexed (WDM) optical burst switching (OBS) is a promising technology for the next generation backbone transport networks. With the increasing use of the Internet to support transport of different traffic types, including that of real-time applications, supporting quality-of-service (QoS) in the optical core network is becoming important. This research focuses on QoS provisioning in WDM OBS networks in terms of service differentiation and fairness. An intrinsic nature of the OBS is the use of offset time where a control packet is sent first to reserve the resources along the route while the data burst is sent after a period of offset time. This feature is important in making high-speed transmission, high data transparency, and all optical switching possible. We explore various issues on QoS provisioning due to the use of offset time as well as developing novel solutions by carefully exploiting this feature.

First, the problem of fast and efficient burst scheduling supporting service differentiation and fairness is considered. Existing scheduling algorithms have either low computational complexity or low burst dropping ratio but not both simultaneously. We propose new algorithms achieving low burst dropping ratio close to the computationally complex algorithm while maintaining the computational complexity at a low level. We develop new burst scheduling techniques called *wavelength reassignment* and *last-hop FDL reassignment* and present new algorithms suitable for classless as well as multi-class environment. These algorithms wisely

make use of the concept of reassignment of the scheduled data burst in the space and (or) time domain before the actual arrival of the data burst; they therefore do not cause any disruption to the on-going traffic. It is important that while providing lower dropping ratio to higher priority traffic, lower priority traffic are not dropped excessively. Our proposed algorithms contribute to the notion of fairness by improving the dropping performance of the lower priority traffic. The performance of the proposed algorithms is evaluated through simulation experiments and the signalling overhead incurred is studied. We show that our proposed burst rescheduling algorithms perform significantly better than existing simple LAUC algorithm in terms of burst dropping probability. At the same time their performance is close to that of the existing complex LAUC-VF algorithm at low loads. The signalling overhead incurred is observed to be less significant when compared to the computational complexity gain achieved over LAUC-VF.

Next, we address the fairness problem in a multi-hop WDM OBS network where different ingress-egress node pairs with different path lengths perform differently within the same class. We develop an efficient fairness method called *link scheduling state based offset selection* (LSOS) with the objective of managing the offset times by choosing offset times based on the link states for bursts with different path lengths such that they perform almost equally. As online link states are used, this method is capable of capturing the traffic loading pattern and topological connectivity. Further, the signalling overhead is low with link state collection done for a short time period only and the offset times computed are used for a sufficiently longer time period. LSOS enables explicit routing with sufficient offset time for node pairs with different hop lengths and under different traffic loading patterns. Further, LSOS is able to achieve fairness with a predefined range of offset time, thus, it ensures that the delay at the edge nodes is at an acceptable level. A simple and efficient scheme, which avoids the need of link states collection done on all the links, avoiding the need for global state information is also presented. We demonstrate the effectiveness of the proposed method

via simulation experiments. We show that the improvement in fairness is achieved with a predefined acceptable range of offset times for classless and multi-class environments with uniform and non-uniform traffic demands.

Finally, we develop a novel scheme for providing edge-to-edge proportional QoS. We propose a *feedback-based offset time selection* (FOTS) method with the aim of providing edge-to-edge proportional dropping ratio among different classes of traffic for various ingress-egress node pairs by dynamically adjusting their offset times. Since the offset time selection is done for the node pairs, FOTS ensures fairness among node pairs with various hop lengths in terms of achieving the proportional QoS. The decision on the use of offset time for various node pairs is done at the edge nodes based on the link states collected by the probe packets. As the intelligent decisions are taken at the edge node rather than the core nodes, FOTS relieves the core nodes of the processing and algorithmic burden. We present an analysis of providing QoS with offset time for a single link model and discuss with numerical results of the analysis, providing the basis for the proposed FOTS method. The effectiveness of FOTS is evaluated through simulation experiments for different values of parameters such as the traffic measurement period, traffic proportion, traffic load, and predefined proportional ratio. We show that FOTS is able to achieve the predefined proportional ratio for node pairs with different hop lengths for various parameters.

Chapter 1

Introduction

With the explosive growth of the Internet as well as various emerging bandwidth-intensive applications such as video-on-demand and video conferencing, the bandwidth demand on the next generation of backbone transport networks will surge in an unprecedented way. Wavelength division multiplexed (WDM) optical networks are a promising candidate for such backbone networks, with hundreds of channels on a fiber each operating at a different optical wavelength [1, 2, 3, 4, 5, 6]. The Internet Protocol (IP) will continue to have a dominant role in communication networks. A straight forward approach to send IP traffic over WDM networks is to use a multi-layered architecture comprising IP-over-ATM-over-SONET-over-WDM. Recently, however, IP-over-WDM networks have received much attention as a promising approach that reduces complexities and overheads associated with the ATM and SONET layers [7, 8, 9, 10, 11].

There are mainly three optical switching techniques that have been proposed in the literature to transport IP traffic over WDM optical networks, namely optical circuit switching (OCS), optical packet switching (OPS) and optical burst switching (OBS). OBS, as described in [12, 13, 14] combines the advantages of OCS and OPS to overcome their shortcomings, thus,

making high data rate, data transparency, and all-optical switching possible.

A major challenge in using WDM OBS networks as the transport infrastructure of the next generation Internet backbone is to provide support for Quality of Service (QoS) differentiation [15]. Mission-critical and real-time applications have more stringent QoS requirements than non real-time applications such as file transfer and email. Much research has been done on supporting QoS differentiation in the Internet with QoS framework such as Integrated Service (IntServ) [16] and Differentiated Services (DiffServ) [17]. However, QoS mechanisms in the Internet such as active queue management and packet scheduling are aided by the availability of electronic buffers at each network node. For the WDM OBS networks, existing optical buffer technologies cannot provide the flexibility and granularity of electronic buffers. Therefore, efficient IP QoS mechanisms are not directly applicable. Instead, now schemes that take into consideration the unique properties of the WDM layer are needed.

1.1 Overview of OBS

OBS is a promising switching technique for the optical Internet since there is no need for buffering and electronic processing of data, which is not the case with OCS. At the same time, like OPS, OBS ensures efficient bandwidth utilization on a fiber link by reserving bandwidth on a link only when data is actually required to be transferred through the link. An OBS burst consists of a control packet (burst header) and a data burst (burst payload) which are sent on separate wavelengths/channels. A data burst is formed by aggregating multiple IP packets at an edge node. The control packet is first sent to reserve the resources along a path and it is followed by the data burst on a separate wavelength after an offset time without waiting for an acknowledgment for the connection establishment. The data burst can pass through the switching nodes along its path all-optically. Since packet processing in the optical domain is

still immature, the control function in the core node still relies on electronic processing. With the burst as a switching unit (rather than an IP packet), the percentage of control overhead as well as the burden on electronic devices in the OBS switches are reduced, thus circumventing the potential electronic processing bottleneck as in WDM OPS¹. OBS takes advantage of the huge capacity in fiber optic transmission systems as well as the sophisticated processing capability in the electronic domain. Not only that OBS can effectively exploit the capabilities of fiber optic transmission systems, it can also facilitate the transition of switching systems in which optical technology plays an important role [13]. OBS is therefore a flexible and feasible solution towards the next generation optical Internet with terabit optical routers and IP over WDM as the core architecture.

A WDM OBS network comprises electronic edge nodes and optical core nodes (OBS switches) interconnected by high-speed WDM links. Each WDM link consists of multiple wavelengths where each wavelength is treated as a channel. An edge node carries out burst assembly/disassembly functions [18]. A core node has an optical switching matrix, a switch control unit and is in charged of forwarding and switching operations. The reader is referred to [14] and [19] for the general architecture and the design of an OBS switch respectively.

The separate transmission and switching of data bursts and control packets can be used to ensure that no buffering of a data burst at intermediate nodes is needed. To realize this, at least δh amount of *offset time* is required, where δ is the control packet processing time and h is the number of hops to be traversed. The control packet processing time includes the time to process the control packet, switching time, time to reserve the appropriate bandwidth, and time to set up the switch [12, 14]. A burst can be optically buffered at a node by using *fiber delay lines* (FDLs). However, FDLs are expensive and hence, is a scarce resource in optical networks. Moreover, they can provide only a very short delay on the order of microseconds.

¹OPS also requires a large number of O-E-O conversion devices to maintain a high data throughput with its higher control overhead per data bit.

Several wavelength reservation mechanisms have been proposed in the literature, e.g., in-band-terminator (IBT), tell-and-go (TAG) [20, 21], and the reserve-a-fixed-duration-based protocol Just-Enough-Time (JET) [22, 23]. These can be distinguished based on how they indicate the end of a burst and the start time of the wavelength allocation. In JET, the burst duration and end time of a reservation are known and the wavelength is open for reservation by other requests after the end time of the current reservation. Therefore, JET with offset time and *delayed reservation* allows statistical multiplexing of data bursts where a wavelength is assigned to a burst for the duration of the burst only. By extending *multi-protocol label switching* (MPLS) capabilities to OBS networks, explicit routing can be used at the ingress nodes [24]. Label switched paths (LSPs) can be set up by sending the signaling messages along pre-determined paths. The control packets and data bursts are then sent along the LSPs. A control packet carries a short label which is swapped at the nodes along its LSP.

Wavelengths are dynamically assigned to bursts. A scheduling algorithm makes the decision in choosing the best wavelength on the outgoing link for the entire transmission duration of the data burst. If no wavelength is immediately available, the data burst is dropped. Several other scheduling algorithms, such as *Latest Available Unscheduled Channel* (LAUC) or *Scheduling Horizon* and *Latest Available Unused Channel with Void Filling* (LAUC-VF), have been proposed in the literature [13, 14, 25]. These algorithms differ in their burst dropping performance and computational complexity.

With the increasing use of the Internet to support the transport of different traffic types, including that of real-time applications, supporting QoS in the optical core network is becoming important where the notion of QoS captures a defined performance contract between the service provider and the end user applications. In general, service differentiation can be provided by specifying various QoS parameters such as delay, burst dropping probability, etc. In a WDM OBS network, the latency of a burst is mainly due to the burst assembly delay at

the edge node, path-setup delay caused by the control packet and the propagation delay in the core network which can be determined. Since OBS uses one-way reservation and bursts are not buffered at the intermediate nodes (if FDLs are used, only a very short delay can be provided), the focus of service differentiation in WDM OBS networks is primarily on the burst dropping performance.

Several methods have been proposed in the literature to support service differentiation in optical networks. As in the Internet, these can be broadly classified into relative and absolute methods. Among the relative differentiation schemes, the extra offset time based method called *prioritized JET* (pJET) [26, 27, 28] assigns an extra offset time to higher priority bursts so that these bursts can make reservations well in advance. This scheme can effectively achieve service differentiation via setting an extra offset time at the edge. It has the advantage that core nodes are relieved of all burdens. However, this method results in long delays and requires large buffers.

The burst segmentation scheme [29] provides service differentiation from a contention resolution perspective. It allows a high-priority burst to preempt a segment of a low-priority burst. Further, burst deflection and composite burst assembly strategies are used. Unlike pJET, the segmentation scheme does not use extra offset times for higher priority classes. However this scheme requires an additional segment header for each segment inside a burst. Also, it incurs extra overhead (signalling message is sent to release the reserved wavelength for the segmented and dropped burst) and increased complexity for burst assembly and reassembly at the edge nodes. More complex scheduling is also needed at the core nodes. Other relative service differentiation schemes include the scheduling based method proposed in [30] and the preemption based methods proposed in [31, 32].

While the above schemes attempt to isolate different classes of bursts, the proportional QoS scheme [33] attempts to maintain the proportion of bursts dropped between different priority

classes by intentionally dropping lower priority bursts. Here, each core node needs to maintain traffic statistics for every individual traffic class. Intentional dropping might result in poor wavelength utilization. Other proportional QoS schemes include the preemptive wavelength reservation scheme in [35, 36], which requires each node to keep track of the usage profile for the respective traffic classes to assist the scheduling decision, i.e., providing proportional QoS via proportional resource allocation [37, 38, 39]. These are basically per-hop based proportional QoS methods and it is not clear how these methods can be extended to support edge-to-edge proportional QoS.

An absolute service differentiation scheme guarantees prespecified dropping probabilities for different classes of bursts. The early drop and wavelength grouping scheme proposed in [40] and preemptive reservation scheme in [41] provide absolute service differentiation through burst admission control and maintaining relevant information at the core nodes.

Another important aspect of QoS support in OBS networks is *fairness*. Fairness in general refers to the requirement that all node pairs belonging to the same class should experience similar performance. Specifically, fairness in an OBS network here refers to requiring, for all ingress and egress node pairs in the network, a burst to have equal likelihood of getting through independent of its hop length to be traversed. It has been observed that node pairs with different hop lengths in an OBS network encounter different burst dropping performance where longer-hop paths perform poorer than shorter-hop paths. A variation of JET called JET-FA has been proposed in [12] to address this issue. The key idea is to assign a fixed extra offset time proportional to the number of hops, allowing a burst on a longer hop path to make resource reservation in advance with its much longer offset time. Again, long delays and large buffers are needed at the ingress nodes. Additionally, shorter-hop bursts tend to be over-penalized. This method is also only applicable to classless traffic and cannot be directly extended to multi-class traffic with varying priorities.

1.2 Motivation and Contribution

In this thesis, we focus on various issues relating to QoS provisioning in WDM OBS networks where the JET protocol and the extra offset time based service differentiation method are adopted. These issues, as shown in Figure 1.1, broadly classified into service differentiation and fairness, include (1) fast and efficient burst scheduling supporting service differentiation and fairness (where the low priority traffic performance is improved significantly at low load), (2) fairness problem due to variation of path length in a WDM OBS network (in a classless as well as a multi-class traffic environment), and (3) providing edge-to-edge proportional QoS to node pairs with various path lengths, thereby ensuring fairness among node pairs with different path lengths.

1.2.1 Fast and Efficient Burst Scheduling

With the enormous bandwidth that a WDM network can offer and an efficient switching technique like the OBS, realizing terabit optical networks as the next generation optical Internet is possible. For supporting such high speed networks efficiently, it is highly desirable that the dynamically arriving bursts are scheduled as quickly as possible. A scheduling algorithm which assigns an available wavelength to a burst for the entire duration of transmission in an efficient way is needed. If fiber delay lines (FDLs) are available, assignment of FDLs to a data burst is required when it cannot be scheduled immediately. The scheduling algorithm has to be computationally simple and has high performance in terms of burst dropping probability.

Due to the dynamic random arrival of bursts with different offset times and hop counts, and the possible use of FDL buffers with varying lengths, a large number of long voids are likely to be created on wavelength channels. Existing scheduling algorithms such as LAUC and LAUC-VF have either low computational complexity or high performance, but not both

simultaneously. This is because LAUC-VF keeps track of the void information and makes use of the voids, while LAUC simply discards the voids.

We develop scheduling algorithms which achieve a balance between the two, that is with high performance close to that of LAUC-VF but with low computational complexity close to that of LAUC. Since the resource reservation decision is made before the actual arrival of the bursts, algorithms which wisely make use of rescheduling techniques to realize the above mentioned goal without causing any disruption to the traffic are introduced. We present two new rescheduling techniques, namely *wavelength reassignment* (reassignment of burst in the space domain) and *last-hop FDL reassignment* (in the space and time domain for network equipped with limited FDLs), to increase the chances of finding a free wavelength for a new burst. We limit the reassignment of bursts in the time domain to the burst traversing the last hop so that no down stream node on the same route will be affected. We develop rescheduling algorithms supporting service differentiation suitable for networks with and without FDLs. It is important that while providing higher burst dropping performance to higher priority traffic, lower priority traffic is not dropped excessively. Our proposed algorithms are attractive since apart from supporting service differentiation, they contribute to the notion of fairness by improving significantly the burst dropping performance of the lower priority traffic at low loads.

1.2.2 Fairness in Multi-Hop WDM OBS Networks

In a multi-hop WDM OBS network, it is important that for all ingress and egress node pairs, a burst has equal likelihood of getting through independent of the hop length it has to traverse. However, bursts traversing longer hop paths have a higher probability of not finding a free wavelength on a link. This results in unfairness. The problem is more pronounced in OBS networks because of the lack of optical buffers at the core nodes, and it occurs in

classless as well as in multi-class environments. Under pJET, a larger offset time is used by a higher priority burst so that its control packet can reserve resources further in advance which increases its chances of finding a free wavelength. However, within a class, bursts with longer hop lengths but without sufficient offset times still experience higher dropping probabilities than those with shorter hop lengths. Existing fairness methods such as JET-FA assign a fixed long extra offset time proportional to the number of hops independent of the network state. This extra offset time is very large, resulting in longer queueing delays and requiring large buffers at the ingress node. Also, this method tends to over-penalize shorter-hop bursts and although applicable to classless traffic, it cannot be directly extended to multi-class traffic.

We develop an efficient fairness method called *link scheduling state based offset selection* (LSOS) with the objective of managing the offset times and to choose offset times based on the link states for bursts with different hop lengths such that they perform almost equally. As online link states are used, this method is capable of capturing the traffic loading pattern and network topological connectivity. Further, the signalling overhead is minimized with link state collection done only for a short time period while the offset times computed are used for a sufficiently longer time period. LSOS enables explicit routing with sufficient offset time for node pairs with different hop lengths and under different traffic loading patterns. Further, LSOS is able to achieve fairness with a predefined range of offset times, thus, it ensures that the delay at the edge nodes is at an acceptable level. A simple and efficient scheme which avoids the link state collection to be done on all the links, thus avoiding the need for global state information is also presented.

1.2.3 Edge-to-Edge Proportional QoS

Recently, proportional QoS, a relative service differentiation model [33, 35, 36], has drawn a lot of attention from the research community due to its ability to provide adjustable performance

spacing between different classes of traffic. Compared to other relative service differentiation models, this model facilitates the pricing process by specifying how well the higher priority traffic will perform relative to the lower priority traffic. Existing proportional QoS models for WDM OBS networks guarantee per-hop proportional QoS via intentional dropping of low priority bursts [33] and preemptive wavelength reservation [35, 36]. The intentional burst dropping method results in poor resource utilization while preemptive wavelength reservation requires the usage profiles of traffic belonging to different classes to be maintained at every node. Also, extra overhead is incurred with the preemptive method. Supporting per-hop proportional loss also does not guarantee edge-to-edge proportional loss [34, 42]. To the best of our knowledge, there is no existing model that has been proposed for providing edge-to-edge proportional QoS in OBS networks.

We propose a *feedback-based offset time selection* (FOTS) method with the aim of providing the edge-to-edge proportional burst dropping ratios between different classes of traffic for various node pairs with different hop lengths by adjusting their respective offset times. Since the offset time selection is done for the node pairs, FOTS ensures fairness among node pairs with various hop lengths in terms of achieving the proportional QoS. The decision on the use of offset time for various node pairs is done at the edge node based on the link state collected by probe packets. As the decisions are taken at the edge nodes rather than at the core nodes, FOTS relieves the core nodes of the processing and algorithmic burdens. As the set of offset times selected is used for a sufficiently longer time before the next probe packet is sent, the signalling overhead is minimized. Further, as online link state information is used, and the offset times needed are computed periodically for supporting edge-to-edge proportional QoS, this method inherently accounts for the traffic loading patterns and network topological connectivity.

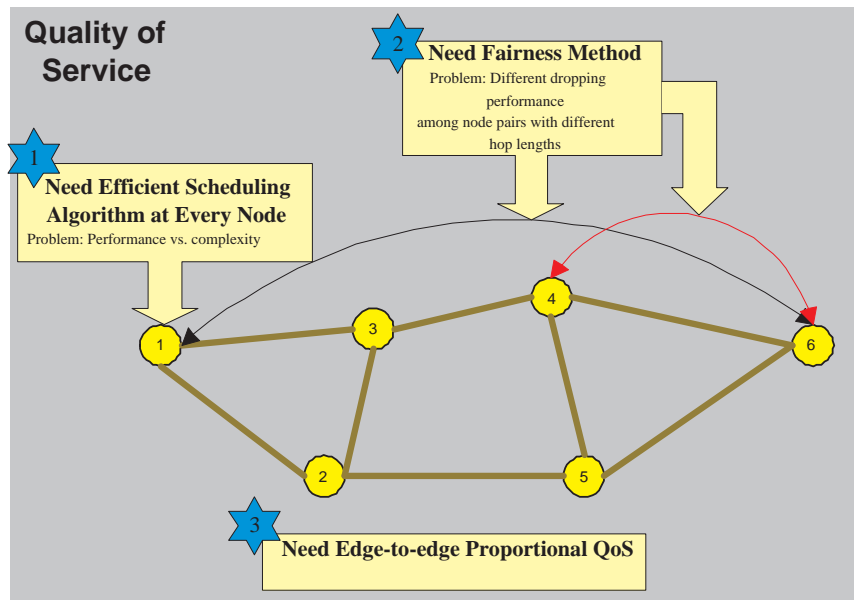


Figure 1.1: Various quality of service issues in WDM OBS networks

1.3 Organization of the Thesis

The rest of the thesis is organized as follows.

In Chapter 2, some background information and related works on WDM OBS networks are presented.

Chapter 3 introduces our novel burst rescheduling techniques and algorithms for WDM OBS networks equipped with and without FDL. We show that the signalling overhead incurred is less significant when compared to the computational complexity gain achieved over existing algorithms. We compare the performance of these algorithms with existing algorithms through simulations.

In Chapter 4, we develop the link scheduling state based offset selection method. The effectiveness of this proposed method is demonstrated for a network with identical and non-identical traffic demands with a predefined range of offset times.

In Chapter 5, we develop the new feedback based offset time selection method. An analysis

of a single link model for the offset time based proportional QoS method is presented. We present numerical results computed from the analytical model to assist the discussion on providing proportional QoS with extra offset time. A simulation performance study is also presented to show its effectiveness.

In Chapter 6, we summarize our research work and discuss possible future extensions.

Chapter 2

Background and Related Work

This chapter discusses the basics of WDM optical networks, optical switching and related work in WDM OBS networks. It focuses on providing the reader with the relevant background information important to this research work. This chapter broadly examines various aspects in WDM OBS networks such as the burst switching protocols, burst scheduling algorithms and service differentiation schemes as well as fairness schemes.

2.1 WDM Optical Networks

WDM optical networks has been a promising choice of solution for today's communication system due to its capability to realize high speed, high bandwidth and improved reliability of service communication channels compared to other existing communication networks. This has been made possible by various enabling technologies for WDM optical networks. Particularly, WDM technology has resulted in increased usable bandwidth without requiring to deploy additional optical fiber. WDM divides optical transmission spectrum into many nonoverlapping channels (wavelength) on a single fiber and allowing every communication

channel to operate at peak electronic speed. WDM optical networks take a few basic forms in terms of network architecture namely the broadcast-and-select networks, wavelength routed networks and linear lightwave networks [3].

The deployment of broadcast-and-select networks is mainly limited to high-speed local area networks (LANs) and metropolitan area networks (MANs) due to the power limitation problem imposed by splitting the transmitted power among various nodes and each nodes receives a fraction amount of the power in the networks. Also, it does not support wavelength reuse as in wavelength routed networks. Wavelength routed networks apart from making better use of the wavelength by allowing wavelength reuse; it does not have the power limitation and scalability problem found in the broadcast-and-select networks. With the introduction of wavelength converters, wavelength continuity constraint can be eliminated. Linear lightwave networks make use of waveband partitioning, where several wavebands are multiplexed on a fiber and several wavelengths are multiplexed on a waveband. By treating waveband instead of individual wavelength as a basic unit, the hardware requirements at the nodes in linear lightwave networks get simplified. In this thesis, we consider the use wavelength convertible (wavelength routed) WDM optical networks.

2.2 Transporting IP Traffic over WDM

Nowadays, most data traffic uses IP, even conventional voice traffic can well make use of voice-over-IP techniques. It is widely believed that IP provides the convergence layer in making the Internet truly ubiquitous [8]. WDM can exploit the use of fiber bandwidth in order to provide enormous bandwidth capacity required for sustaining the continuous growth in the Internet traffic. Hence, it has emerged as a core transmission technology for the next generation Internet backbone networks. There are three main approaches for sending IP

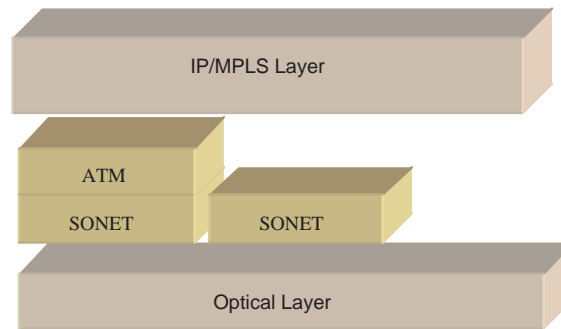


Figure 2.1: Possible protocol stack options for IP-over-WDM

traffic over WDM as shown in Figure 2.1. The first one is to transport IP-over-ATM-over-SONET/SDH-over-WDM. The second and third approaches are IP-over-SONET/SDH over-WDM and IP-over-WDM, respectively. ATM provides QoS and traffic engineering support and SONET provides the protection/restoration capability which is transparent to the upper layers such as the IP layer. However, such benefits are offset by the substantial overheads needed, (for example SONET carries overhead information which is encoded in several levels), inefficiency in usage of bandwidth for data-centric IP applications using fixed bandwidth allocation (SONET) or fixed-size cells (ATM), as well as the complexity to manage and control (for example managing IP over ATM compared to IP-leased line network) the network [43]. Among these, IP-over-WDM is the most efficient solution as it reduces the overheads and complexities associated with the ATM and SONET layers.

2.3 Optical Switching Techniques

There are several switching methods to transfer IP traffic over WDM networks such as optical circuit switching (OCS), optical packet switching (OPS) and optical burst switching (OBS). The following sections briefly discuss the above mentioned burst switching techniques.

2.3.1 Optical Circuit Switching

With OCS, there are three distinct phases which are the circuit (lightpath) setup, message transfer and circuit release phases. Dedicated lightpaths need to be set up (including configuring the switches along the paths and receiving the acknowledgement at the source sent from the destination) before data is transferred. OCS does not use statistical multiplexing and hence, does not make use of the resources efficiently especially for bursty Internet traffic.

2.3.2 Optical Packet Switching

OPS [84] allows IP traffic to be processed and switched on a per packet basis at every router in the network. An IP packet has two parts called the header and payload. The header carries the necessary information such as source and destination node IP addresses and is sent together with the data packet along the same path. Upon reaching a router, the header packet is processed electronically (including forwarding and switching) and the data packet is optically buffered using FDLs. The switching of an optical packet has been evolving from conventional packet switching in the electronic domain to switching in the optical domain to increase the switching speed. Apart from expediting the packet switching, an OPS network supports statistical multiplexing and hence utilizes the network resources more efficiently compared to OCS. However, synchronization of packets, switching hardware cost, and other technological limiting factors are preventing OPS from becoming commercially viable in the near term.

2.3.3 Optical Burst Switching

A burst in OBS has two parts which are referred to as the *control packet* and the *data burst*, respectively. Unlike OPS, OBS decouples the control and data as shown in Figure 2.2. The

control packet and data burst are sent over separate channels known as the control channel and data channel, respectively, through the OBS network. This way, OBS makes use of the sophisticated control in the electronics domain with control packet processed electronically while the data burst is switched optically. A burst is a super packet assembled at an ingress router by aggregating a number of IP packets destined to the same egress router with similar requirements, e.g., QoS. This allows the switching overhead to be amortized across many packets. The concept of temporal separation between the control packet and the data burst in OBS has made it possible to bypass the need for buffers to account for the delay incurred by the processing of the control packet at each and every intermediate node. This is done by having the control packet sent first to reserve the required wavelength for the upcoming data burst while the data burst is stored for a long enough amount of time at the ingress node before being transmitted into the network so that it will never overtake the control packet. Note that at the ingress node (edge router) electronic buffers are abundant whereas within OBS networks, optical buffer is a scarce resource with very limited delay functionality. In a later section, we will describe various OBS techniques to facilitate better understanding of how the above advantages are realized in an OBS network.

2.4 OBS Networks

A typical OBS network is shown in Figure 2.3. The network comprises edge nodes and core nodes. These nodes are connected by high-speed WDM links. A WDM link carries multiple wavelengths (say W wavelengths) where each wavelength is treated as a channel. With W channels on one link, there are W_c control channels (e.g. wavelength w_0 in Figure 2.2) and $W - W_c$ data channels (e.g. wavelengths w_1 and w_2 in Figure 2.2) for the transmission of control packets and data bursts, respectively. An edge node carries out burst assembly [18]/dis-assembly functions and provides legacy interfaces. Burst assembly is carried out at

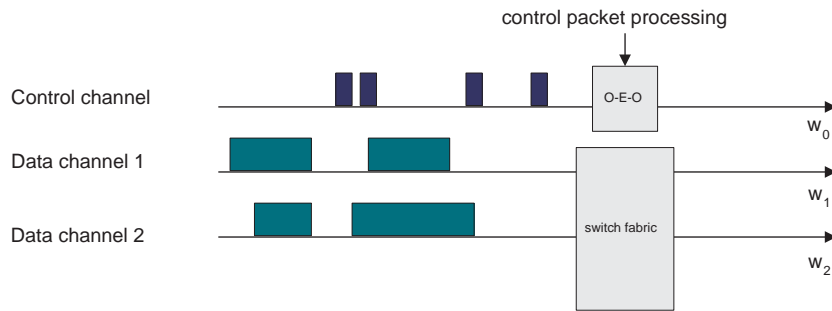


Figure 2.2: Separation of control channel(s) and data channel(s) in OBS.

an ingress node. Basic burst assembly schemes are either timer-based or threshold-based [44]¹. A control packet is sent first followed by its corresponding data burst after a time gap. Both traverse a number of core nodes before reaching the destined egress node. A core node has an optical switching matrix and a switch control unit (SCU), and is in charged of forwarding and switching operations. Upon reaching the egress node, the data burst is disassembled into IP packets which are then transmitted to the respective access networks. Figure 2.4 shows the general architecture of an OBS node [14]. An OBS node has N input fibers and M output fibers, with each fiber carrying W wavelengths. It uses N number of $W \times 1$ wavelength demultiplexers and M number of $1 \times W$ multiplexers. Each fiber has a data channel group (DCG) of $W - W_c$ channels and a control channel group (CCG) of W_c channels. The SCU has functionality similar to a conventional electronic router. The input FDLs if available can be used to delay the data bursts so that the SCU has enough time to process its associated control packets. The optical buffers of FDLs are used to resolve contention on the outgoing data channels. Routing and control protocols are run on the routing and signalling processors.

¹Recently, there are several research works done on burst assembly such as [19, 45, 46, 47] which focus on burst assembly algorithm, the effect of burst assembly on the performance of the OBS network.

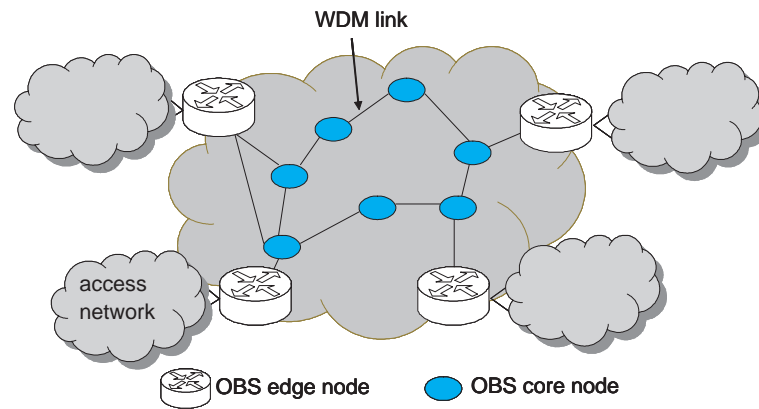


Figure 2.3: An optical burst switching network.

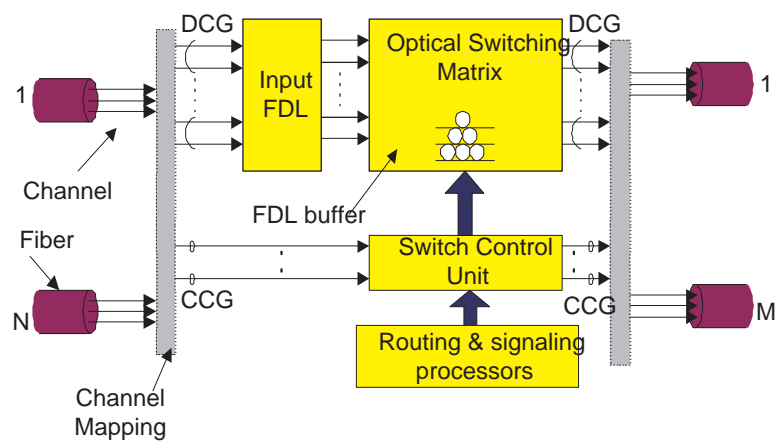


Figure 2.4: General architecture of an OBS node

2.5 Optical Burst Switching Techniques

An efficient optical burst switching technique supports bursty traffic in an on-demand manner, ensuring that high bandwidth utilization as well as high burst dropping performance are achieved. With OBS, the control packet precedes the data burst with an amount of time referred to as the offset time, denoted by $T_{off} \geq 0$, while the data burst is stored at the ingress node in the electronic domain for the period of T_{off} . The control packet reserves bandwidth for the corresponding data burst as it traverses along a path. While the control packet is processed at each intermediate node, the data burst will cut through the pre-configured node all optically if the reservation has been successful. If no wavelength is available, the request is blocked and the corresponding data burst is dropped. In case of congestion or output port conflicts, the burst is dropped as well.

There are several variations of burst switching techniques proposed in the literature such as in-band-terminator (IBT), tell-and-go (TAG) [20, 21], just-in-time (JIT) [22, 48, 49, 50], and reserve-a-fixed-duration-based (RFD-based) protocol just-enough-time (JET) [22]. They differ in the way bandwidth is reserved/released and the choice of T_{off} . These protocols and several variants also differ in other ways such as the hardware requirements, the signaling architecture (e.g., in-band or out-of-band), performance, complexity and cost. Despite their differences, the common feature of all these OBS protocols is that they all involve one-way reservation which greatly reduces the pre-transmission delay of a burst. Note that this is important as the burst transmission time can be relatively short given the high speed links.

In IBT, each burst has a header and a special delimiter which is used to indicate the end of the burst. The special feature of IBT virtual cut-through [51], which allows a source and any intermediate node to transmit the head of a burst even before the tail of the burst is received. Since no store-and-forward operation is needed with virtual cut-through, there is little delay for a burst. However, the release of the reserved wavelength for reservation from the other

bursts is upon detection of the delimiter. The reserved wavelength is therefore not open for future reservation from other bursts.

In TAG, the control packet is first sent on a separate control channel to reserve wavelength along a path for the following data burst². The data burst is transmitted on the data channel after some offset time T_{off} . Similar to circuit switching, a control signal is sent to release the wavelength after the burst is sent. It is different from circuit-switching in that no acknowledgment is needed in order to send the data burst out.

In the JIT protocol (equivalent to the TAG scheme) [48], the data burst is sent after some offset time T_{off} , but it reserves a wavelength immediately upon processing the control packet at a node. This is referred to as *immediate reservation* (IM). As shown in Figure 2.5, at node i the wavelength is reserved starting from t' , the time at which the control packet has been processed. The burst however will arrive at a later time t_a . Since the control packet is not aware of the burst length, the end of the transmission is not known to the node until an explicit release message is sent to release the bandwidth or a time out occurs. The shaded region in Figure 2.5 represents the wavelength reservation period.

With JIT, the network node keeps track of whether the wavelength channel is currently reserved or not. Once the channel is occupied, the channel status is set to RESERVED, and any new control packet arriving at a node that sees a RESERVED status will not be able to use that channel. Upon receiving the explicit RELEASE message or when it times out, the status of the channel will be updated to FREE. Any control packet arriving at a node and sees a FREE status will be allowed to reserve the channel. JIT is therefore conceptually simple. However, apart from inefficiency due to its open-ended wavelength reservation, JIT does not

²The purpose of sending the control packet first includes informing each intermediate node of the upcoming data burst, configuring the switch fabric (so that the burst to be switched to the appropriate output port) and making the routing decision.

make use of the wavelength effectively with its IM. This is because a burst will not arrive at the node before the offset period is reached but the channel is reserved upon processing of the control packet, resulting in an idle period from t' to t_a . Despite all these, JIT has been implemented and deployed in the OBS field trial [52] since it is significantly simpler than the JET protocol (to be described later)³.

RFD burst switching differs from TAG by reserving the wavelength for a specified amount of time. This can be implemented by specifying in the control packet the offset time and the burst length. Making use of this information, the RFD-based OBS protocol called JET has two unique characteristics which are the use of an offset time and *delayed reservation* (DR). The control packet is sent first and the data burst is stored at the edge node for at least some *basic offset time* δh (to ensure that the data burst will not overtake the control packet), where δ is the processing time of the control packet (that includes switch set up time and control packet transmission time) at each node and h is the hop length. The control packet contains such information as the destination address, data burst length and offset time for the corresponding burst arrival. The offset time in the control packet is adjusted for the next hop upon successful reservation. For example, a 3-hop burst has the basic offset time of 3δ at the ingress node (first node) and the offset time on the control packet is adjusted to 2δ at the second node and δ at the third node if the control packet has been successful in wavelength reservation on these links.

Given a node i as shown in Figure 2.6, DR makes a reservation on the chosen wavelength from time t_a (after an offset time T_{off}), the time at which the burst is expected to arrive instead of the time t' at which the control packet has been processed. The wavelength is reserved until the burst departure time, $t_a + l_b$, where l_b is the expected burst length (the shaded region in Figure 2.6 is the wavelength reservation period). JET, which uses extra information to

³For a detailed description, evaluation, and comparison of the JET, JIT, and etc, the reader is referred to [53]

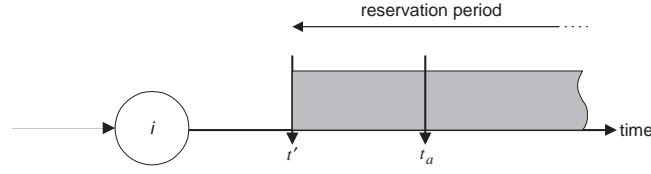


Figure 2.5: The use of offset time and immediate reservation in JIT.

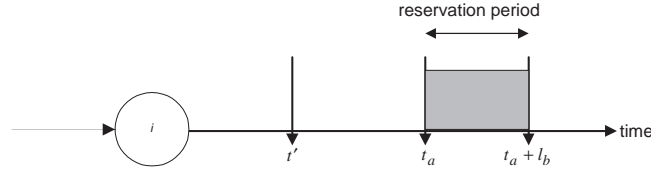


Figure 2.6: The use of offset time and delayed reservation in JET.

better predict the start time and end time of the burst, thus facilitates more efficient resource utilization of the bandwidth and buffer (FDLs, if used) than IBT-based and TAG-based burst switching techniques.

2.6 MPLS Framework for IP-over-WDM

We now discuss briefly how MPLS [10] can be used to enable explicit routing at the ingress node of an OBS network. In contrast to the traditional destination-based hop-by-hop forwarding in IP networks, MPLS uses labels in making forwarding decisions at the network nodes. *Forward equivalence classes* (FECs) are used to represent the possible forwarding options. Bursts destined for a given egress node and with the same service requirement may belong to the same FEC. The bursts are initially labeled at the ingress node depending on the FEC to which they belong. At an intermediate node, a new label is used to determine the next hop for the burst. The incoming label is replaced with the outgoing label which identifies the respective FEC for the downstream node. Such a label based forwarding method reduces the processing overhead involved in routing at the intermediate nodes. With MPLS, *explicit*

routing can be done at the ingress node to set up the *label switched path* (LSP). The explicit route is carried by the burst only at the time of set up. Once the LSP is set up, a burst traversing a specific path is forwarded by using a label⁴. LSPs can be seen as semi-permanent data pipes between various ingress-egress nodes which can be set up and torn down and where subwavelength allocation is permitted (since bursts still reserve resources based on the optical burst switching technique used). Such features allow MPLS to play an important role in expediting the transfer of traffic through the network and at the same time ensuring efficient utilization of network resources.

2.7 Scheduling Algorithms

In OBS networks, a burst is transmitted without any acknowledgement of successful reservation of the resources along its path. Therefore, a burst can be dropped at any intermediate node if its control packet fails to reserve a free wavelength. Subsequently, all the resources successfully reserved along the path for this burst at the upstream nodes will be wasted. Therefore, a scheduling algorithm plays a vital role to yield improved burst dropping performance as well as good resource utilization in an OBS network.

The role of a scheduling algorithm is to assign an available wavelength to a burst for the entire duration of its transmission in an efficient way. If fiber delay lines (FDLs) are available, assignment of FDLs to a data burst is required when it cannot be immediately scheduled upon the arrival of the burst. If no resource is available, the control packet is blocked and the corresponding data burst will be dropped. If the scheduling algorithm is not able to find a suitable wavelength fast enough before the arrival of the burst, the burst will also be dropped.

⁴A control packet thus contains information such as offset time, length of burst and label (routing information).

In the literature, several methods have been proposed to resolve burst contention. Burst segmentation has been proposed as an effective contention resolution technique [54]. Burst scheduling algorithm using burst segmentation and FDLs has been proposed in [55] to handle data burst contention and reduce burst dropping probability compared to existing scheduling algorithm. An analytical framework for the performance study of networks using burst segmentation has been presented in [56]. Other research work on contention resolution proposed in the literature are such as deflection routing (space deflection) [57] and look-ahead window contention resolution [58]. In [59], the shortcomings of the existing contention resolution schemes have been highlighted especially when the offered load is excessively high. It has proposed using a modified TCP decoupling approach to control the offered load to the OBS switch. The performance of TCP has been studied in terms of delay and throughput with various OBS networks characteristics and parameters in [60].

Since the arrival of a burst at a node in OBS is dynamic, a burst is assigned to a wavelength in real time on a per burst basis. Therefore, time on the wavelength is fragmented into periods where the wavelength is occupied and idle (void). Further, the use of offset times of varying lengths also contributes to creating the voids. Voids are idle period on the wavelength which result in wastage of bandwidth resources. Few existing scheduling algorithms try to maximize the wavelength utilization and burst dropping performance by minimizing the voids created by new bursts and scheduled bursts. It is preferred to assign a burst to a channel (wavelength) that will become available just shortly before the burst arrives. This way, the void created is minimized. Further, by doing so, the burst dropping probability is reduced as a new burst is always tightly packed with existing bursts, leaving more space for future bursts.

Scheduling algorithms such as *Latest Available Unscheduled Channel* (LAUC) or *Scheduling Horizon* and *Latest Available Unused Channel with Void Filling* (LAUC-VF), have been proposed in the literature [13, 14]. They differ in terms of their computational complexity and

burst dropping performance. The computational complexity includes how much information needs to be kept at a scheduler and how complex (in terms of time) the scheduling operations are.

While LAUC and LAUC-VF aim to maximize the overall dropping performance of the network by maximizing the performance locally on all the outgoing links connected to a node, there are also some other proactive scheduling algorithms available in the literature taking into consideration burst dropping at downstream nodes when making a scheduling decision on the first outgoing link. These proactive algorithms include the *Priority-based Wavelength Assignment* (PWA) algorithm [61] and the *Burst Overlap Reduction Algorithm* (BORA) [62]. These two scheduling algorithms aim to reduce possible burst contention at downstream nodes and therefore improve the overall burst dropping performance.

2.7.1 LAUC

In LAUC scheduling, for every wavelength, only one *unscheduled time* on the wavelength, is maintained. The unscheduled time is the time after which no reservation has been made (the open end segment). Upon the arrival of a burst, LAUC examines the unscheduled time at every wavelength and selects the wavelength with the latest available unscheduled time (refer to Figure 2.7, W_4 is selected for the new burst as it is the latest available one out of two available channels, i.e., W_1 and W_4) which can accommodate the burst at the requested time. By choosing the latest available channel, the void formed by the new burst with the previous burst scheduled on the same channel will be minimized. Once a wavelength is selected, the burst is scheduled on this wavelength and the unscheduled time for this wavelength will be updated to the end time of the new burst. If there are K wavelengths, the time complexity of this algorithm is $O(K)$, which is reasonably low. However, since the void intervals are not used, LAUC results in poor bandwidth utilization and high burst dropping probability.

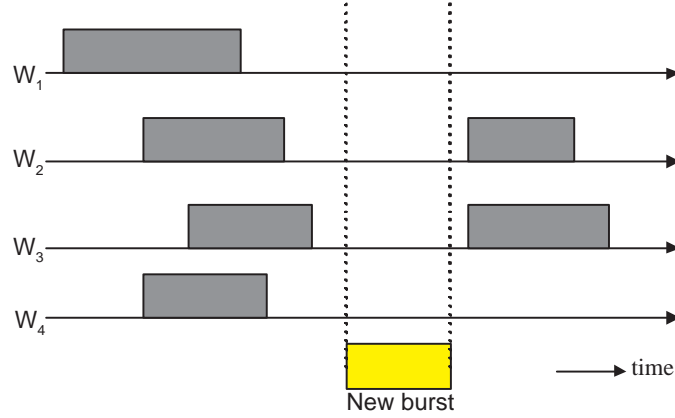


Figure 2.7: Illustration of LAUC and LAUC-VF.

2.7.2 LAUC-VF

LAUC-VF maintains all the void intervals information on every wavelength⁵ and allows a newly-arrived burst to be scheduled in the *unused time*. Unused time is the time in between the scheduled bursts (i.e., the void) or the open end of the last burst on the wavelength. LAUC-VF first checks the unused times on all the channels and chooses the latest available one for the duration of the burst at the requested time. For example, in Figure 2.7, for LAUC-VF, W_1 through W_4 are available, however, W_3 is selected as it is the latest available unused channel. Intuitively, LAUC-VF is computationally more complex than LAUC as it maintains information on all the voids and examines through the voids when scheduling takes place. If there are a maximum of V voids on one of the K wavelengths, the time complexity of LAUC-VF is $O(KV)$ if searched linearly [14]. Despite its high computational complexity, LAUC-VF has better bandwidth utilization and burst dropping performance compared to LAUC. In [63], the minimum starting void (Min-SV) algorithm uses an augmented binary search tree to simplify the search for a suitable wavelength. The time complexity of this algorithm is reduced to $O(\log KV)$ compared to $O(KV)$. An algorithm similar to LAUC-VF was introduced in [64] with the difference of minimizing the ending void instead of minimizing

⁵Start time and end time of scheduled bursts are maintained.

the starting void in LAUC-VF.

2.7.3 PWA

PWA [61] is used at the ingress node which keeps a wavelength priority database for every destination node. The wavelength priority is dynamically updated according to its burst dropping profile (feedback from the network). In other words, the priority reflects the likelihood of the burst transmission on the wavelength to be successful (considering dropping due to wavelength contention at the intermediate node). When a new burst arrives, it will check for availability on the wavelength with the highest priority first, followed by the next highest if the highest is not free to accommodate the burst at the required time. PWA can reduce burst dropping probability in an OBS network compared to random assignment with the aid of continuous monitoring of the burst dropping performance. However, it is only applicable in an OBS network without wavelength conversion capability.

2.7.4 BORA

BORA [62] also tries to reduce burst dropping at downstream nodes due to contention by scheduling a burst proactively at the edge node. This is done with the aid of electronic buffers at the edge nodes for reducing the overlapping degree⁶. A higher overlapping degree can result in higher contention and hence a higher dropping probability. Under this scheme, the burst which has a portion of time overlapped with another burst will be delayed (bounded) although there is actually a free wavelength to be used without any delay. This is achieved by searching the wavelength in the ordered manner called *fixed order search*, and the algorithm stops either when a suitable channel is found which satisfies the maximum delay requirement

⁶Overlapping degree is referred to the number of bursts that arrive at one link simultaneously for a given time.

or when the list is exhausted. This action, referred to as serializing (placing side by side) the burst, aims to reduce burst dropping by reducing the overlapping degree on the first link on each OBS path. Another *destination-based order search* scheme [62] is introduced and is shown to reduce burst dropping further by taking the routing information of different OBS paths into consideration while scheduling the locally generated bursts. Under this scheme, every burst has a *home channel* (note that the term channel here refers to the path and not the wavelength) that will be searched first. If the burst fails to be accommodated by the home channel, the scheduler will search other paths. Each home channel has a different preferred wavelength. This action causes two bursts traversing the same path to be serialized on the same wavelength (not overlapped in time) and therefore reduces possible contention in the intermediate nodes. Among the several algorithms (collectively known as BORA) proposed to reduce the overlapping degree, some make use of void and some simply discard the void while searching for a suitable wavelength. Compared to PWA, BORA does not rely on the feedback from the network (hence the need to collect loss information from the network). However, BORA requires more complex processing at an ingress node.

2.8 QoS Provisioning

With the growing popularity of various emerging real-time Internet applications such as Internet telephony, video on demand and video conferencing, the shift of the network service model from best-effort to quality aware is anticipated. To realize such networks, scalable, efficient and robust QoS schemes are needed, considering the tremendous amount of traffic that the backbone networks have to handle as well as the limitations (e.g., FDL scarcity) imposed by the WDM layer.

QoS in general refers to the capability of a network to provide better service to selected

network traffic. For example, there can be two types of traffic denoted by class 1 and class 2, where class 1 traffic might belong to email applications and class 2 traffic might belong to those of real time applications⁷. A network which is capable of providing QoS will be able to provide better service to class 2 traffic where its stringent requirement demands a higher QoS than the class 1 traffic.

Extensive research has been done on providing service differentiation in the switching nodes of conventional Internet Protocol (IP) networks. A key mechanism to provide service differentiation is active queue management which operates to enqueue packets to meet different packet loss probabilities. Service disciplines govern the scheduling and buffering at a switching node and hence control the interaction between different classes and differentiate the traffic [65]. In an OBS network, a data burst is assembled at an ingress edge node by aggregating the packets of the same QoS class and with the same destination egress node. Bursts from different connections requiring different QoS will interact with each other at a switch, which without proper control, the performance of the network could be affected. Although various queue management disciplines have been proposed and their performance have been widely studied, they are not directly applicable to the OBS network since they mandate the use of buffers. In OBS networks, the use of electronic buffer is not desirable (O-E-O conversion causes loss of data transparency and increases the network cost and delay) while optical buffers being a limited resource in an optical network, provide only a very limited delay functionality. Further, per-hop based methods do not guarantee edge-to-edge performance in a network environment. In view of these, developing efficient QoS methods taking into account the unique properties of the WDM layer is needed in order to make future optical Internet QoS capable. Further, computationally simple solutions are needed for suitable implementation at very high speeds.

Existing service differentiation methods in OBS networks are broadly classified into relative

⁷Throughout this thesis, class 1 refers to low priority traffic and class 2 refers to high priority traffic are used.

and absolute differentiation methods. A relative service method differentiates the QoS received by different classes in a relative manner, i.e. qualitatively. However, an absolute method provides a quantified QoS. Relative differentiation methods include the extra offset time based method called pJET [28], the burst segmentation scheme [29], the scheduling based method [30], the preemption based method [31], the probabilistic preemption method [32], the per-hop proportional QoS by intentional dropping method [33] and preemptive wavelength reservation [35, 36]. Absolute QoS differentiation is considered by the early drop and wavelength grouping scheme in [40] and the preemptive reservation scheme in [41].

2.8.1 Offset-time based Service Differentiation

The offset time, apart from facilitating the transmission of bursts optically through the core nodes (refer to the description of JET in Section 2.5), can also be adjusted to support QoS. The concept of using extra offset time to achieve class isolation in an OBS network was first proposed by Yoo *et al.* in [28, 66] in their prioritized JET protocol called pJET. Consider two classes of bursts. Class 1 belongs to non real-time applications such as email. Class 2 belongs to higher priority service which is used for delay sensitive applications such as real-time audio and video communications. Here, class 1 traffic can afford to be retransmitted but not class 2 traffic due to stringent time constraints. Therefore, it is important that the burst dropping probability of class 2 traffic be lower than class 1 traffic. The key idea of pJET is to assign a fixed *extra offset time* on top of the basic offset time to bursts from class 2 traffic, but only a basic offset time to class 1 traffic. This results in advance reservation of class 2 traffic, giving better chance of success than class 1 traffic.

The goal of pJET is to achieve complete class isolation by assigning a long extra offset time to higher priority bursts. For example, a higher priority burst belonging to class m is isolated from the lower priority burst of class $m - 1$ by assigning to it an extra offset time of $5L_{m-1}$ (for

99% class isolation) where L_{m-1} is the mean burst length for the lower priority burst. Since the offset time is fixed at the edge node, no extra computation is needed at the core nodes hence, this service differentiation scheme is simple. However, this method incurs longer delays and hence requires large buffers at the ingress nodes. For the higher priority traffic, although the pretransmission delay is longer, it is balanced out by a lower dropping probability which in turn reduces delays caused by retransmissions in the upper protocol layers.

Yoo *et al.* considered burst length distribution of the lower priority traffic to determine the extra offset time needed to achieve complete class isolation. Barakat *et al.* [67] considered other factors which could affect the extra offset time needed in this context. This includes the ratio of the arrival rates of lower priority traffic and higher priority traffic. They showed that increasing the ratio of the lower priority traffic to the higher priority traffic increased the extra offset time needed for complete isolation. This suggested that achieving complete class isolation through assigning a fixed amount of extra offset time (regardless of the traffic's arrival rates) as proposed by Yoo *et al.* might not be adequate especially when the lower priority traffic was dominant.

The simplicity of this offset time based method and its suitability to be implemented in bufferless WDM networks have generated much interest in the method. Various analytical models have been introduced to approximate the burst dropping performance. Particularly, [28] has provided analysis on determining the upper and lower bounds of the burst dropping probability, queuing delay and degree of isolation using the offset time. Dolzer *et al.* [68] and So *et al.* [69] analyzed the burst dropping probability in an OBS node taking into consideration the contention between low and high priority traffic in a two-class case while Kim *et al.* [70] analyzed the blocking probability for general n classes ($n > 2$), i.e., a multi class system. There are also other analytical models introduced such as [71] by Barakat *et al.* which is suitable for quantifying the burst dropping probability of a multi class non-work

conserving OBS system⁸.

2.8.2 Segmentation based Service Differentiation

The burst segmentation scheme [29] achieves service differentiation in the perspective of contention resolution in the optical core nodes. This is done by preferentially segmenting and deflecting bursts of different priorities when resolving contentions in the core. For example, when contention occurs, a high-priority burst is allowed to preempt a segment of the low-priority burst. Burst segmentation can be implemented with deflection. With deflection, a burst can be deflected entirely or just partly (if the burst has been segmented). As a result, the burst or a segment of the burst is not dropped but deflected, thus increasing the chances of the burst reaching its destination. Alternatively, packets of different priorities can be assembled into different segments inside a burst in decreasing priority from head to tail (referred to as a composite burst^{9,10}). When contention occurs, the low priority segment of the contending burst is preempted. Different policies on whether to deflect/ segment/ drop or some combination of these are illustrated for different scenarios; the reader is referred to [29] for details. The achievable factor of improvement with burst segmentation has been analyzed in [73], where the limit of the ratio (i.e., the limit of possible improvement with burst segmentation) between the blocking probabilities of burst segmentation and JET policies on a single optical burst switch is shown. However, with burst segmentation, the control packet, updated with the new burst length information, has to be sent to release the unnecessary wavelength (duration) already reserved for the burst (before segmentation). This results in

⁸The assumption of a work-conserving OBS system is used in most of the above models.

⁹The performance of optical composite burst switching (OCBS) has been compared to OBS by Neuts *et al.* in [72] assuming Poisson-distributed traffic. Also, Detti *et al.* [54] developed an analytical model for OCBS with an ON-OFF arrival process.

¹⁰The dropping probability is calculated in terms of packets (of upper layers) rather than the OBS burst.

extra signalling overhead which can be excessive if segmentation takes place frequently. Also, additional segment headers for the segments inside a burst are necessary. Burst assembly at the ingress node and burst reassembly at the egress node are also more complex. Complex scheduling is also needed at the core nodes. In view of all these, careful study on this scheme is needed in order to make it an effective solution.

2.8.3 Scheduling based Service Differentiation

A scheduling algorithm to support QoS in the OBS networks has been proposed in [30]. This scheme partitions data bursts into n classes depending on their QoS requirements. At a core node, data bursts that use the same output link are scheduled by the scheduler associated with that link where n queues are maintained. These queues are served one by one in priority order by using LAUC-VF. This scheduling algorithm therefore generalizes the LAUC-VF scheduling algorithm to include Diffserv QoS features in a straightforward manner. It has been shown that performance improvement has been achieved for one class at the expense of the other. Also, buffers are needed at every node within the network. A framework for realizing IP Diffserv over OBS networks has been presented in [74].

2.8.4 Preemption based Service Differentiation

The preemption based scheme proposed in [31] provides strict priority for high priority traffic by dropping scheduled bursts belonging to lower priority traffic using a preemptive scheduling algorithm. The probabilistic preemptive scheme proposed in [32] allows a high priority burst to preempt a scheduled low priority burst in a probabilistic manner when there is no free wavelength to accommodate the new burst. Since the preemptive probability p affects the burst dropping probability for the high priority traffic as well as the low priority traffic, p can

be adjusted to get different burst dropping probability ratios. Analysis of this scheme with a multi-dimensional Markov chain has been presented.

2.8.5 Proportional Service Differentiation

While the above schemes attempt to isolate different classes of bursts, the proportional QoS model attempts to maintain the proportion of burst dropping between different priority classes, for example $\frac{d_1}{d_2}$, where d_1 and d_2 are the dropping probabilities of class 1 traffic and class 2 traffic, respectively. Network operators are able to adjust the service differentiation spacing between classes by adjusting in proportion the class differentiation parameters, for example $\frac{s_1}{s_2}$, where s_1 and s_2 are the differentiation parameters of class 1 and class 2 traffic, respectively. The focus of proportional QoS is therefore on how to make sure the ratio of the burst dropping probabilities obeys the predefined ratio of the class differentiation parameters (i.e. ensuring $\frac{d_1}{d_2} = \frac{s_1}{s_2}$). In [33], the proportion is maintained by intentionally dropping lower priority bursts when the predefined ratio of the class differentiation parameters are violated. Maintaining different statistics for every individual traffic class is needed at every core node. Further, under this scheme, bursts can be dropped even if there is an idle channel. As a result, this approach results in a higher overall blocking probability and hence poorer network utilization.

In [35], a preemptive scheduling technique has been proposed to provide proportional QoS in terms of burst loss and data channel usage. The preemptive wavelength reservation method requires each node to keep track of the usage profile for the respective traffic classes to assist the scheduling decision. That is, to provide proportional QoS via proportional resource allocation [37, 38, 39]. A profile (either usage or loss depending on the performance criteria intended for service differentiation) for each class is maintained at a node. A class of traffic is said to be in profile if its current usage does not exceed the predefined usage limit (given by the

service differentiation parameters) and is out of profile otherwise. The proportional service differentiation is achieved with the use of partial preemptive scheduling. For example, in the case of achieving proportional data channel usage, when scheduling takes place for a new burst, if a free wavelength is available, the burst is scheduled and the profile is updated. If there is no free wavelength available, the service differentiation profile is checked and compared to the preset service differentiation parameters. If it is out of profile, the algorithm then searches and selects a wavelength which has a scheduled burst with the smallest overlapping part with the new burst. The overlapping part of the new burst is dropped while the rest is scheduled on the selected wavelength. Otherwise, if in-profile, the algorithm searches for a eligible class of scheduled burst (which is out of profile) by examining their respective profiles and identifying the burst to be preempted partially (which has the smallest overlapping) or entirely so that the new burst is scheduled in its entirety. If no burst is eligible for preemption, the new burst will be dropped partially or entirely.

In [36], a similar method is proposed by Liao *et al.* to support proportional QoS in such a way that the basic concept of using preemptive scheduling including that of partial preemption and full preemption is used. The authors suggested that blocking probability at a node was inversely proportional to the number of idle wavelengths, and their proposed method controlled the fractions of resources different classes were allowed to use. Different from [35], this scheme did not allow a new burst to be partially scheduled. An analytical model was derived to quantify the dropping probability and utilization of the wavelength. This method was shown to alleviate the low bandwidth utilization problem in [33] where proportional QoS was achieved via intentional burst dropping.

2.8.6 Absolute Service Differentiation

Different from relative QoS methods which specify QoS qualitatively, absolute service differentiation guarantees dropping probabilities for different classes of bursts to be no worse than the specified values (i.e., quantitatively). Existing absolute service differentiation methods for OBS networks require information to be maintained at every node, and effective burst admission control is needed. The early drop and wavelength grouping schemes proposed in [40] provide absolute service differentiation through burst admission control and also management of wavelengths. The early drop scheme selectively drops non-guaranteed traffic in a probabilistic manner (where the early dropping probability is decided based on the online measured burst loss probability at every node and the predefined maximum per-hop loss probability) while wavelength grouping manages the wavelengths for guaranteed traffic. The early dropping of lower priority bursts is intended to avoid contention with bursts of higher priority. Wavelength grouping mechanism on the other hand supports absolute loss differentiation at each OBS node. This is done by classifying traffic into groups and each group is assigned with a unique label and a provisioned set of wavelengths. It has been shown that the integration of these two mechanisms is effective in providing the worst-case dropping probability to the guaranteed traffic and reduces significantly the dropping of the non-guaranteed traffic.

The preemptive reservation scheme in [41] provides edge-to-edge loss probability through ensuring per-node absolute loss guarantees. In this scheme, each node allocates a loss threshold to every QoS class (which is the upper bound on loss probability of the respective class). Making use of preemptive differentiation and an admission control mechanism, the thresholds are maintained at the per-node level. The edge-to-edge loss probability is then divided into per-node loss probabilities and are allocated to the intermediate core nodes. By maintaining the per-node absolute loss guarantees, the edge-to-edge loss probability can be delivered. The preemptive differentiation scheme makes use of the loss threshold to select a contending burst

to be dropped. The scheme selects the burst with the largest distance to the threshold in the contention list constructed upon contention (i.e., when a control burst arrives at a node and fails to reserve an outgoing wavelength), thus ensuring that all the distances to thresholds belonging to different classes at every node are kept equal. When the offered load increases, the distance to threshold of all classes at a node will converge to zero, thereby making sure that no individual flow can be breaching its threshold and affecting those classes that conform to their thresholds. With equal distance to threshold for every class, the admission control mechanism makes sure that the average of these is greater than zero in order to maintain the loss thresholds.

2.9 Fairness

Generally speaking, fairness is the ability of a network to provide the same level of service to all its users within a class. In a WDM OBS network, fairness is an important issue where edge-to-edge service guarantee should be maintained for ingress-egress node pairs with different path lengths. For example for all ingress and egress node pairs in a network, a burst should have equal likelihood of getting through independent of the hop length it traverses. At a finer level, fairness also refers to bursts of different sizes having equal likelihood to reach their destinations. The burst length fairness problem is obvious in an OBS network with LAUC-VF wavelength reservation algorithm [75] where longer bursts are dropped more frequently than shorter bursts. This is because fragmented voids are usually small. As a result, bursts from a specific ingress router experience consistently poorer dropping performance if they employ burst assembly mechanisms which result in longer bursts.

The path length fairness problem has been addressed in [12]. It has been observed that bursts traversing various paths with different hop lengths experience different burst dropping per-

formance. Specifically, bursts traversing longer hop paths have higher dropping probabilities compared to those bursts traversing shorter hop paths due to the higher chances of encountering a bottleneck link. JET-FA [12] has been proposed to solve this problem in a classless environment. This is done by assigning extra offset time proportional to the hop length, e.g., $5hL$ or $10hL$, where h is the hop length and L is the mean burst duration of the lower priority burst. There are several drawbacks associated with this approach. As the maximum hop length becomes large, the extra offset time will be large and therefore a longer queuing delay is needed at the ingress node. Also, this extra offset time is fixed and does not take into consideration the network state such as the traffic load. This method is developed for a classless traffic environment and cannot be directly extended for a multi-class environment.

2.10 Summary

In this chapter, we have presented the basics in optical switching as well as various related works in WDM OBS networks. Various optical switching choices are described first while the remaining discussion has been focused on the preferred OBS switching technique. The related works surveyed included the few variations of optical burst switching techniques, burst scheduling algorithms, service differentiation schemes as well as fairness schemes. We have also highlighted the special features as well as the drawbacks of the various schemes.

Chapter 3

Burst Rescheduling Algorithms

In this chapter, we consider the problem of fast and efficient dynamic scheduling of bursts that belong to different classes of priority in WDM-based OBS networks for 2 switch types; without any buffers and with limited optical buffers. In OBS networks, control and data components of a burst are sent separately with a time gap to ensure that resources such as wavelengths (and FDLs if provided) are reserved at various nodes before the data burst arrives. A *scheduling algorithm* with attractive features such as computational simplicity and efficient resource utilization is mandatory to quickly handle dynamic burst traffic and reduce burst dropping probability. Existing scheduling algorithms such as LAUC and LAUC-VF are either computationally simple or have good burst dropping performance but not both simultaneously as explained in Chapter 2. Further, there may arise situations wherein the above algorithms fail to schedule a new data burst to some wavelength due to the non-availability of resources. Since scheduling is done well in advance before a data burst actually arrives, any changes to the allocated bursts are possible. The idea of burst rescheduling which assigns a scheduled data burst to other available wavelength to accommodate a new data burst is therefore a way to improve burst dropping performance. Also, in a multi-class environment where extra offset time is assigned to the higher priority traffic for advance reservation, the

high dropping performance of the high priority traffic comes with the expense of the poor dropping performance of low priority traffic, since high priority traffic with its large assigned offset time always win those with shorter offset time in terms of wavelength reservation. The idea of rescheduling increases the chances of finding a free wavelength for the low priority traffic thereby improving its dropping performance. We propose *burst rescheduling* as an alternative to void filling which can do fast scheduling without requiring to examine and fill voids and at the same time can achieve good dropping performance.

The focus of this chapter is to develop new burst scheduling algorithms which could realize effective optical burst switching for dynamically arriving requests (carried by the control packets) with high performance close to that of LAUC-VF but with low computational complexity close to that of LAUC. Based on the idea of rescheduling, two algorithms namely, *On-Demand Burst Rescheduling* (ODBR) and *Aggressive Burst Rescheduling* (ABR) are developed. Further, we develop an algorithm called Burst Rescheduling with Wavelength and Last-hop FDL Reassignment (BR-WFR) which is suitable for networks equipped with limited FDLs. These algorithms support classless as well as multi-class traffic environment. We discuss the complexity issue, signalling overhead and feasibility of implementing burst rescheduling. Through simulation experiments we demonstrate the effectiveness of the proposed burst rescheduling algorithm.

The rest of the chapter is organized as follows. In Section 3.1, the details of the proposed burst rescheduling techniques are presented followed by the burst rescheduling approaches in Section 3.2. We present the proposed burst rescheduling algorithms and discuss the complexity issue in Section 3.3. The signalling overhead and the feasibility of implementation are discussed in Section 3.4 and Section 3.5 respectively. Section 3.6 presents the results of performance study while Section 3.7 summarizes the work presented in this chapter.

3.1 Burst Rescheduling Techniques

As both low computational complexity and low burst dropping probability do not coexist in the two existing scheduling algorithms, namely LAUC and LAUC-VF, we propose new scheduling algorithms which combine the relative merits of both algorithms (high performance with low dropping and low complexity) with the motivation that a scheduled burst can be rescheduled to another available wavelength in order to accommodate a new request. This is possible as requests arrive dynamically and a control packet reserves wavelengths well before the arrival of its corresponding data burst. There is a variable time gap between the arrival time of data burst and that of control packet. This time gap, which is also referred to as the offset time, depends on the total number of hops that need to be traversed for the data burst from the source to the destination. With this variable offset time between the control and data burst, there arise situations wherein a burst could be possibly rescheduled to another available wavelength before it arrives. It is to be noted that rescheduling does not affect any ongoing traffic. Rescheduling a burst on a link requires changes in the control setting in both the end nodes of the link. Therefore, whenever rescheduling is successful, a special “NOTIFY” control packet is sent to the next node to notify about the changes, for e.g., wavelength for that burst so that the receiving node will do the necessary settings.

The benefit of rescheduling is illustrated in Figure 3.1. As shown in Figure 3.1(a), there are two wavelength channels (or wavelengths), W_1 and W_2 on a given fiber. Bursts 1, 2, 3, and 4 are assumed to arrive one by one in the given order. As the request of burst 1 (i.e. control packet of burst 1) arrives first, wavelength channel W_1 is assigned to it. Request of burst 2 arrives next and it is scheduled to channel W_2 because W_1 is not available. The request of burst 3 arrives next. Since W_1 and W_2 are both free, the latest available unscheduled channel to minimize the gap formed, in this case, W_1 , is chosen. Burst 4 is allocated to W_2 since W_1 is not available (still occupied by burst 1). Burst 5, with the duration of L arrives

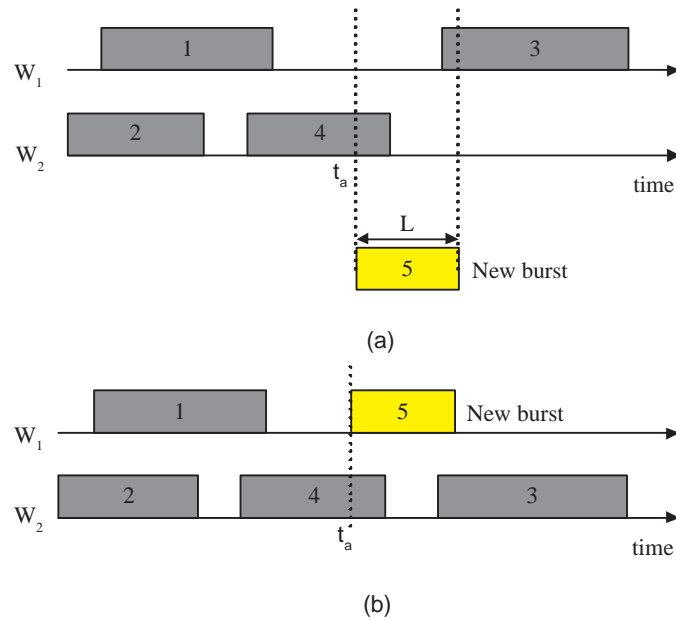


Figure 3.1: Illustration of the benefit of burst rescheduling. (a) Both LAUC and LAUC-VF fail to schedule the new burst. (b) The new burst is scheduled by rescheduling burst 3.

at time t_a which requests for the time slot as shown in Figure 3.1(a), will be dropped by LAUC or LAUC-VF due to the non-availability of wavelengths. As shown in Figure 3.1(b), by rescheduling burst 3 from W_1 to W_2 , data burst 5 could be allocated to W_1 .

When OBS nodes are provided with FDLs, burst rescheduling can be done by using wavelength reassignment method and or the Last-hop FDL reassignment method. These methods are illustrated below.

3.1.1 Wavelength Reassignment

With wavelength reassignment, the rescheduling process changes only the wavelength assigned to a burst keeping the scheduled time unchanged. Wavelength reassignment considers reassigning a scheduled burst (last burst on a wavelength) from a wavelength to another available wavelength at the open end (therefore, after wavelength reassignment, the burst is still the last burst on the new wavelength). Figure 3.2 illustrates the benefit of wavelength reassignment.

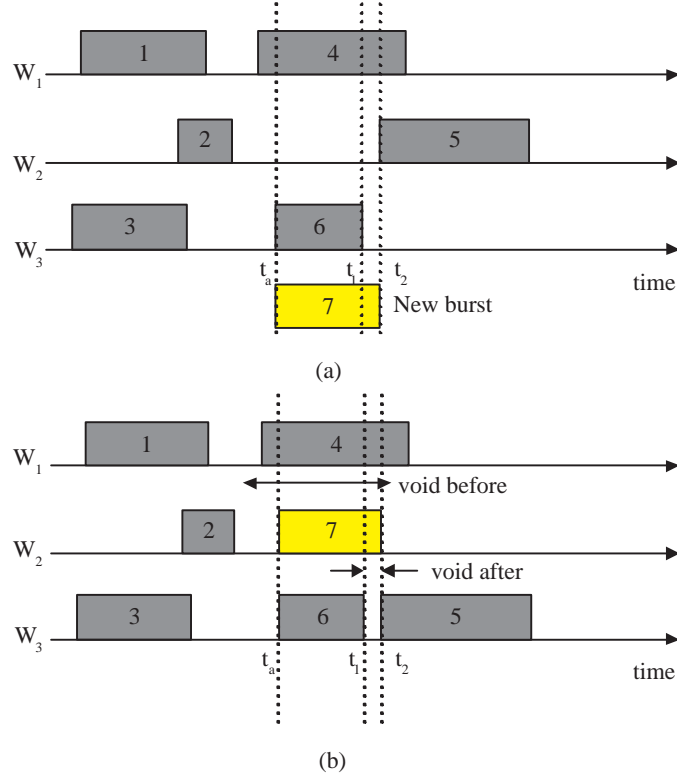


Figure 3.2: Illustration of the benefit of wavelength reassignment. (a) LAUC fails to schedule burst 7. (b) Burst 7 can be scheduled by using wavelength reassignment.

Figure 3.2(a) shows that without wavelength reassignment, burst 7 cannot be scheduled. As burst 1 through burst 6 have been allocated one by one to wavelength channels W_1 , W_2 or W_3 based on the latest available wavelength similar to LAUC, burst 7 arriving at time t_a cannot be scheduled to any wavelength. Figure 3.2(b) shows that with wavelength reassignment, burst 7 can be scheduled successfully. Here, if burst 5 at t_2 is reassigned from W_2 to W_3 (with a shorter void $(t_2 - t_1)$ formed by burst 5 with burst 6 at W_3 than with burst 2 at W_2), when burst 7 arrives, it can be allocated to W_2 .

3.1.2 Last-hop FDL Reassignment

The idea is to reassign an FDL to change the schedule of a burst which traverses its last hop. We do not consider the FDL reassignment at hops other than the last-hop because

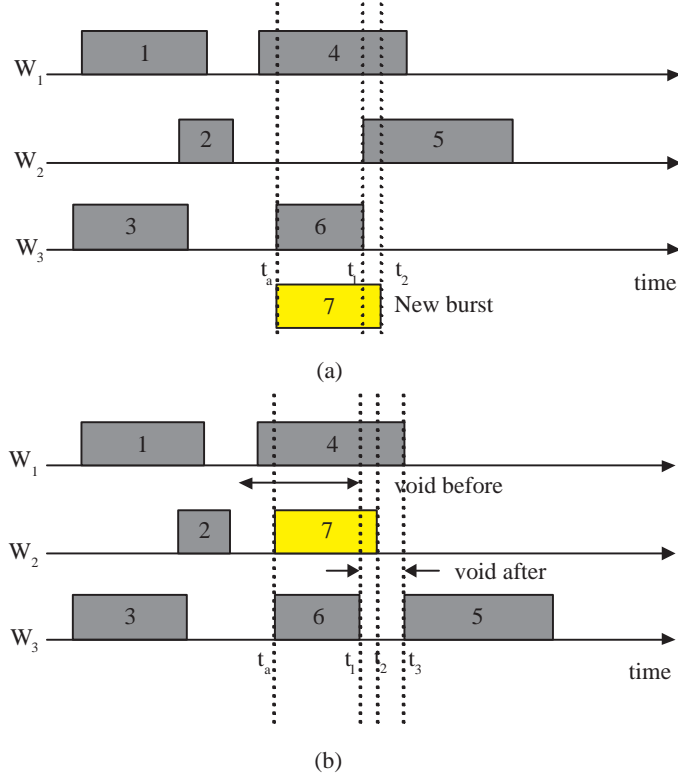


Figure 3.3: Illustration of the benefit of burst rescheduling with FDL reassignment. (a) LAUC fails to schedule the new burst, wavelength reassignment does not help. (b) The new burst is scheduled by allowing FDL reassignment.

(i) FDL reassignment changes the burst arrival time at the downstream nodes and there is no guarantee that the required resources will be available, (ii) processing and signalling overhead could become high as every downstream node need to be notified. However, last-hop reassignment does not pose these problems. The benefit of FDL reassignment is illustrated in Figure 3.3. In Figure 3.3(a) burst 1 through burst 6 are scheduled to the latest available wavelength in the given order. With this, burst 7 arriving at time t_a cannot be scheduled. Figure 3.3(b) shows that burst 7 can be scheduled if FDL reassignment is used. When burst 6 is successfully scheduled neither burst 4 nor burst 5 can be rescheduled to the current wavelength W_3 , i.e., no wavelength reassignment can take place. However, if burst 5 can be delayed by using a free FDL, it can be rescheduled to wavelength W_3 (the void formed at W_3 is smaller than void at W_2) and burst 7 can be scheduled to wavelength W_2 .

3.2 Burst Rescheduling Approaches

Rescheduling algorithms can be developed based on two approaches. They are known as *single-level rescheduling* and *multi-level rescheduling*. Single-level rescheduling involves only one burst to be rescheduled to another available wavelength to accommodate a new burst. The examples illustrated in Figure 3.1(b), 3.2(b), and 3.3(b) fall under this category. In multi-level burst rescheduling, several bursts are rescheduled one by one in sequence to other available wavelengths in order to accommodate a new burst. As shown in Figure 3.4(a), no wavelength is available at time t_a for the new burst if single-level rescheduling is used. Multi-level rescheduling can reschedule burst 4 from W_2 to W_3 followed by rescheduling burst 2 from W_1 to W_2 to free wavelength W_1 to accommodate the new burst, as shown in Figure 3.4(b). Multi-level rescheduling is expected to provide better dropping performance than single-level rescheduling. However, from computational complexity point of view, multi-level rescheduling is more complex than single-level rescheduling. This is because a multi-level rescheduling algorithm needs to determine an appropriate order (among several possibilities) in which different bursts are to be rescheduled in sequence.

Since our goal is to develop faster algorithms to schedule dynamically arriving bursts, the proposed burst rescheduling algorithms are based on single-level rescheduling approach.

3.3 Burst Rescheduling Algorithms

Having discussed various burst rescheduling techniques and approaches, we now develop the burst rescheduling algorithms. First, we introduce ODBR and ABR which make use of wavelength reassignment. Further, we develop BF-WFR which make use of one or both of - wavelength reassignment and last-hop FDL reassignment which is suitable for WDM OBS networks equipped with limited FDLs.

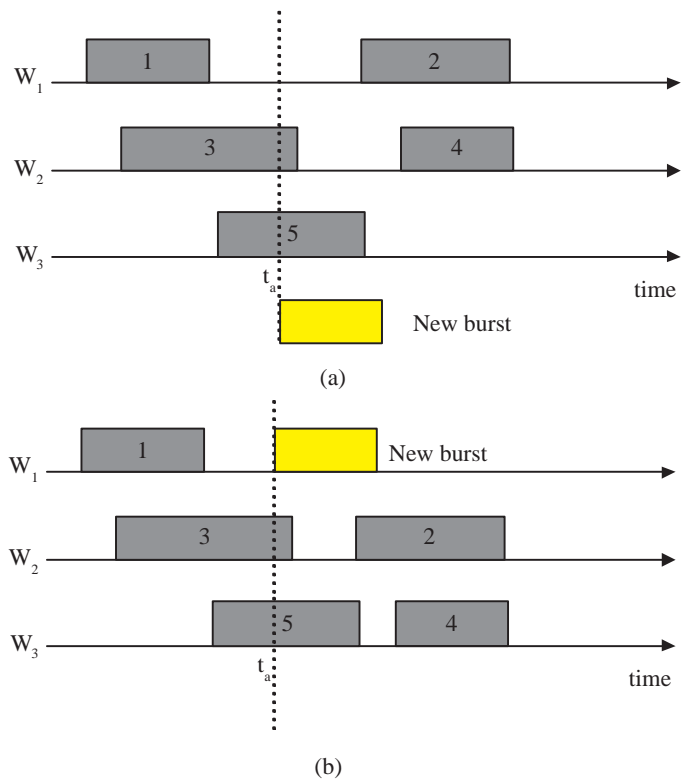


Figure 3.4: Illustration of multi-level rescheduling. (a) No wavelength is available for new burst. (b) Rescheduling of burst 4 from W_2 to W_3 followed by rescheduling of burst 2 from W_1 to W_2 frees W_1 to accommodate new burst.

3.3.1 On-Demand Burst Rescheduling (ODBR) Algorithm

As the name suggests On-Demand Burst Rescheduling (ODBR) algorithm considers rescheduling of an existing burst only if it is necessary. ODBR is invoked when a burst fails to be scheduled to any of the wavelengths. The goal of ODBR is to reschedule a scheduled burst to another available wavelength so that the wavelength is available for the new burst. The algorithm works in two phases. When a new request arrives phase 1 is executed to select a suitable free wavelength similar to LAUC. If no wavelength is available, phase 2 is called to check if any of the existing bursts can be moved to a new wavelength to enable scheduling of the new burst. The algorithm examines the wavelengths one by one. For a given wavelength, it checks if the last burst can be moved to any other wavelength and determines the void created. After examining all the wavelengths, it chooses the one which possibly creates the smallest void after migration.

The pseudo-code of the algorithm is given in Table 3.1. The new burst is assumed to arrive at time t . The latest available time of wavelength channel W_i is denoted by t_i .

A simple example shown in Figure 3.5(a) and (b) helps to illustrate how ODBR works. Let the arrival time of the new burst be t . Phase 1 fails to assign any wavelength to the new burst as shown in Figure 3.5(a) and phase 2 will therefore be invoked. Wavelength W_1 and W_3 both have a valid out-wavelength W_2 . Therefore, rescheduling of last burst from W_1 or W_3 to W_2 would make a wavelength available for scheduling the new burst. In order to optimize the performance, ODBR chooses the best wavelength which has the latest available time. In this case, the void formed by the new burst on W_3 by rescheduling the burst from W_3 to W_2 is the smallest compared to the void formed at W_1 by rescheduling the burst from W_1 to W_2 . Therefore, W_3 is the best wavelength. The last burst on W_3 is rescheduled to W_2 and the new burst can be scheduled to W_3 as shown in Figure 3.5(b).

Table 3.1: ODBR algorithm

Phase 1

For every wavelength W_i compute m_i as $t - t_i$ if $t_i \leq t$; otherwise set m_i to infinity. Choose wavelength W_j such that m_j is finite and is minimum among all. If no such W_j exists, call phase 2; otherwise assign wavelength W_j to the new burst.

Phase 2

Step 1: For every wavelength W_i , determine if out-wavelength V_i is valid. Out-wavelength is valid if the last burst on W_i can be moved to V_i and the new burst can be scheduled to W_i , otherwise out-wavelength is said to be invalid for W_i .

Step 2: If no valid out-wavelength V_i exists, the new burst is dropped. Otherwise, choose wavelength W_p such that V_p is valid and is the latest available wavelength after rescheduling among all the valid out-wavelengths.

Step 3: Reschedule last burst on W_p to V_p . Assign new burst to W_p .

Step 4: Send a special “NOTIFY” control packet to notify the next node about the change in wavelength of the rescheduled burst.

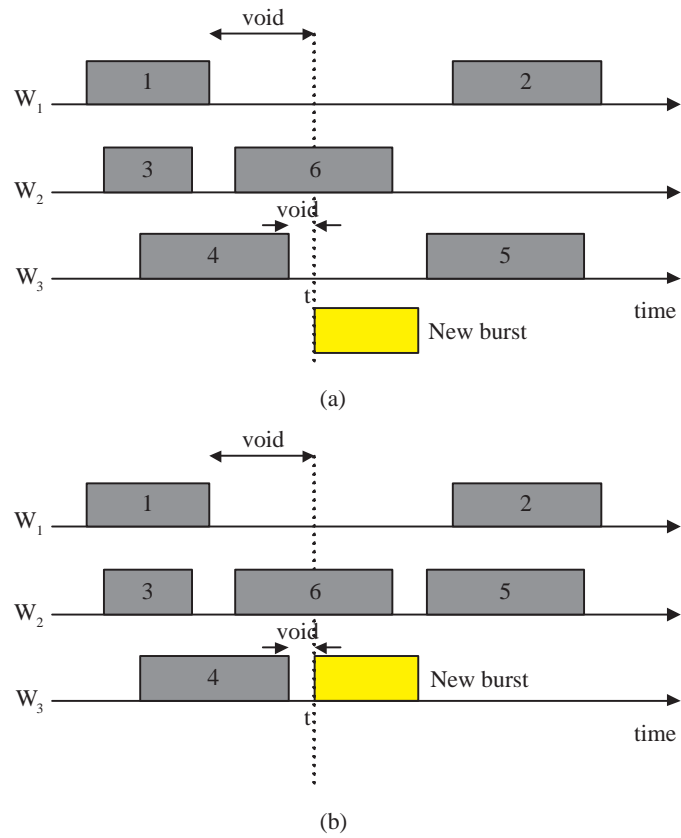


Figure 3.5: Illustration of ODBR. (a) A situation wherein the new burst can not be scheduled. (b) The last burst on W_3 is moved to W_2 to accommodate the new burst on W_3 .

Complexity of ODBR

- Phase 1 - Scheduling : ODBR examines the information of one burst on each wavelength. Phase 1 runs in $O(W)$ time in the worst case.
- Phase 2 - Rescheduling : Phase 2 has the worst case complexity of $O(W^2)$ time since it examines the last burst on each wavelength for rescheduling to one of the other wavelengths.

Since the complexity of LAUC-VF is $O(KW)$, the complexity of ODBR will be more than LAUC-VF if $W > K$. However, since ODBR is called only when a burst is dropped (usually less than 10%), the overall processing complexity remains better than LAUC-VF. It therefore has the advantage of low complexity similar to LAUC. We note that the algorithm ABR described in the next section has the worst case complexity of $O(W)$ for phase 1 and phase 2.

3.3.2 Aggressive Burst Rescheduling (ABR) Algorithm

As shown in Table 3.2, the ABR algorithm has two phases as well but it is different from the ODBR algorithm in that phase 2 is not invoked when phase 1 fails but when phase 1 is successful. This algorithm is intended to prevent future data burst dropping by invoking rescheduling every time a burst has been scheduled successfully. In ABR, upon successful scheduling of a burst at W_p in phase 1, rescheduling of one latest burst from some other wavelength W_i to W_p takes place in phase 2 if such a burst exists. Rescheduling is governed by the rule that the void formed when the burst is rescheduled from W_i to W_p is minimum among all possible wavelengths. By doing so, the probability of dropping data bursts that arrive later could be decreased.

Examples as shown in Figure 3.6 and Figure 3.7 are used to illustrate this algorithm. Figure 3.6 shows two wavelengths and bursts that are being considered. Burst 1, 2, 3, 4, and 5

Table 3.2: ABR algorithm

Phase 1

For every wavelength W_i compute m_i as $t - t_i$ if $t_i \leq t$; otherwise set m_i to infinity. Choose wavelength W_p such that m_p is finite and is minimum among all. If no such W_p exists, drop the burst; otherwise assign wavelength W_p to the new burst and call phase 2.

Phase 2

Step 1: For every wavelength W_i other than W_p determine if the last burst can be rescheduled to W_p and also the void created at W_p after rescheduling. If such rescheduling is possible for a wavelength, it is said to be a valid in-wavelength for W_p . If no valid in-wavelength exists phase 2 fails.

Step 2: Choose a valid in-wavelength W_j which has the smallest void.

Step 3: Reschedule the last burst from W_j to W_p .

Step 4: Send a special “NOTIFY” control packet to notify the next node about the change of wavelength of the rescheduled burst.

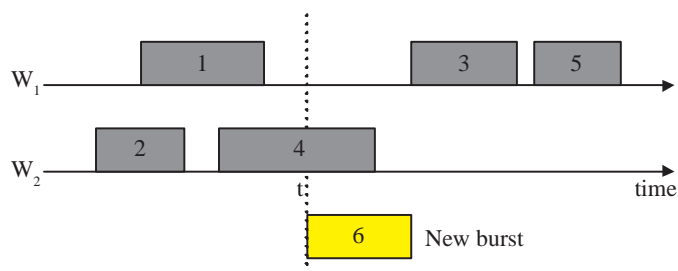


Figure 3.6: Illustration of a situation wherein LAUC, ODBR and LAUC-VF fail to schedule new burst 6.

arrive at a node one by one in that order and are scheduled to W_1 and W_2 at phase 1. When burst 6 arrives, it could not be scheduled to any wavelength by LAUC, LAUC-VF, or ODBR. For the same burst arriving pattern, Figure 3.7(a) to (d) show that with ABR, the new burst 6 which would otherwise have been dropped, could be scheduled successfully. This demonstrates that prevention of burst dropping could be achieved by using ABR. As shown in Figure 3.7(a), when burst 4 is scheduled to wavelength W_2 , consideration for rescheduling of one last burst from other wavelength to current wavelength W_2 takes place. Here, the last burst on wavelength W_1 is scheduled to wavelength W_2 as it conforms to the rule that void formed after rescheduling is shorter, as shown in Figure 3.7(b). Figure 3.7(c) shows that when burst 5 arrives, it is scheduled to wavelength W_2 in phase 1 as it is the latest available wavelength. Finally, burst 6 will be scheduled to W_1 at time t of its arrival as shown in Figure 3.7(d). If there are more than one burst on different wavelengths that could be rescheduled to W_p , the burst with the smallest void formed after being rescheduled to W_p would be chosen. This is to make sure that the smallest void would be formed every time a rescheduling takes place. It is worth noting that without ABR, the situation as shown in Figure 3.6 could not be handled by LAUC, ODBR or even LAUC-VF.

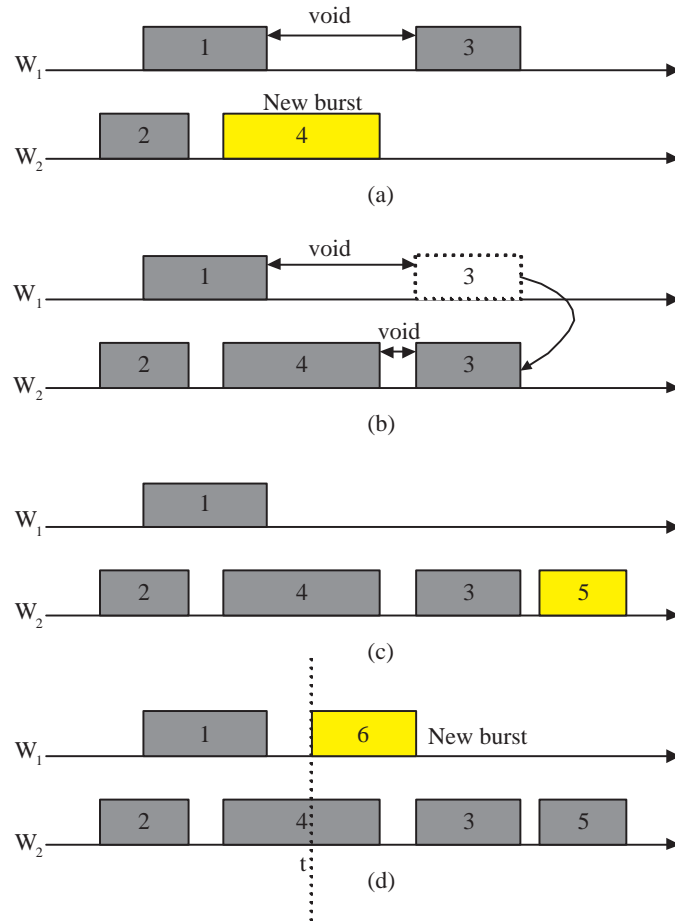


Figure 3.7: Illustration of working of ABR. (a) New burst 4 is assigned to W_2 . (b) Last burst from W_1 is rescheduled to W_2 . (c) Burst 5 is assigned to W_2 . (d) Burst 6 will be able to be scheduled to W_1 .

Complexity of ABR

- Phase 1 - Scheduling : ABR examines the information of one burst on each wavelength. Phase 1 runs in $O(W)$ time in the worst case.
- Phase 2 - Rescheduling : Phase 2 has the worst case complexity of $O(W)$ time as well since it examines only the last burst on each wavelength for rescheduling.

The complexity of ABR is approximately two times that of LAUC since it examines only the last burst on each wavelength for rescheduling. Therefore, even for values of $K \geq 3$, ABR is expected to run faster than LAUC-VF whose complexity is $O(KW)$. It therefore has the advantage of low complexity similar to LAUC.

3.3.3 Burst Rescheduling with Wavelength and Last-hop FDL Reassignment (BR-WFR) Algorithm

Since ABR has better advantage in terms of complexity compared to ODBR, we develop BR-WFR for WDM OBS networks equipped with limited FDLs with rescheduling done in the aggressive way rather than on-demand. With BR-WFR, whenever a burst is successfully scheduled, burst rescheduling uses one of both of the rescheduling techniques - wavelength reassignment and last-hop FDL reassignment - on an possible existing burst to help increase the chances of accepting bursts that arrive later. Consider the case where every node is equipped with a limited FDL buffer of size F with FDL_i ($i = 1, 2, \dots, F$) capable of optically delaying data by i time units, the BR-WFR algorithm has two phases called scheduling and rescheduling phases. Phase 1 examines only the last burst on every wavelength to find the best wavelength W_p (latest available) and an FDL (if needed). Phase 2 takes place if phase 1 is successful and an existing last burst from $W_j (\neq W_p)$ is considered to be rescheduled to W_p to make a shorter void in order to pack the bursts tightly to prevent future burst dropping

(wavelength reassignment equivalent to ABR). Phase 2 may also reassign an FDL used by a burst if wavelength reassignment is not possible and only if the outgoing link is the last hop for the burst. Burst rescheduling with wavelength reassignment (BR-WR) in which only wavelength reassignment is carried out, is a special case of BR-WFR. The pseudo-code of the algorithm is given in Table 3.3. Here, the new burst is assumed to arrive at time t . The latest available time of wavelength channel W_i is given by t_i .

Complexity of BR-WFR

- Phase 1 - Scheduling : BR-WFR examines the information of one burst on each wavelength. Also, it searches through the FDLs once. Therefore, the worst case complexity is $O(W + F)$.
- Phase 2 - Rescheduling : This involves searching through all the wavelengths once for the best last burst to be rescheduled and also the FDLs once, if delay is needed. Therefore it requires $O(W + F)$ time. BR-WFR therefore has the same computational complexity as LAUC which also runs in $O(W + F)$ time.

Thus, the overall computational complexity of burst rescheduling at a node is given by $O(W + F)$.

3.4 Signalling Overhead

Additional signalling is needed when rescheduling is successful by using ODBR, ABR or BR-WFR algorithm. This is to notify the next node about the change of wavelength by sending a “NOTIFY” packet. However, rescheduling does not incur significant signalling overhead for both ODBR, ABR, and BR-WFR algorithms as illustrated in the numerical examples below

Table 3.3: BR-WFR algorithm

Phase 1 [Burst Scheduling]

Step 1 [Wavelength search]: For every wavelength W_i compute the gap, m_i as $|t - t_i|$. If $t_i \leq t$, set the FDL flag to false otherwise set the flag to true to indicate that an FDL is needed.

Step 2 [Wavelength assignment]: Choose wavelength W_p with false FDL flag (i.e., no FDL needed) such that m_p is minimum among all m_i . Assign wavelength W_p to the new burst and call phase 2; if no such W_p exists, go to step 3.

Step 3 [FDL assignment]: Choose W_p which has minimum m_p . Search for shortest FDL which is at least m_p long. Assign this FDL and wavelength W_p to the new burst and call phase 2. If no such W_p exists, drop the burst and exit.

Phase 2 [Burst Rescheduling]

Step 1 [Wavelength search]: For every wavelength W_i other than W_p determine if the last burst can be rescheduled to W_p ; compute the void V_i that would be created at W_p after rescheduling; set the FDL flag to true if the burst needs to be delayed.

Step 2 [Wavelength reassignment]: Choose wavelength W_j with false FDL flag such that V_j is the minimum among all the wavelengths. If no such W_j exists, go to Step 4.

Step 3 [Notification]: Reschedule the last burst from W_j to W_p . Send a special control packet “NOTIFY” to notify the next node about the change of wavelength of the rescheduled burst. Exit.

Step 4 [FDL reassignment]: Choose W_k whose last burst is at the last hop and V_k is the minimum among all the wavelengths. Search for the shortest FDL which is at least V_k in length. If no such W_k exists, rescheduling fails; exit.

Step 5 [Notification]: Reschedule the last burst from W_k to W_p . Send a special control packet “NOTIFY” to notify the next node about the change of wavelength and the change of arrival time.

and it has also been verified through simulation experiments (will be reported in respective performance study sections). Out of all the bursts that invoke the rescheduling process, only a fraction of these are successful in rescheduling. It is worth noting that “NOTIFY” control packet is much smaller than the “RESERVE” control packet as it needs to carry only the wavelength change information (only two fields with wavelengths W_j and W_p). Also, no complex algorithm is executed upon receiving the “NOTIFY” packet. Further, a successful reschedule requires a “NOTIFY” packet to be sent on only one link. Alternatively, without sending extra signalling packet, these information can be piggybacked to the control packet. It therefore does not incur significant processing time and does not consume significant control channel bandwidth when compared to the computational complexity gain achieved over existing algorithms such as LAUC-VF.

3.4.1 Signalling Overhead for ODBR

To observe the signalling overhead incurred by ODBR, consider a network node where x of the total bursts arrived have been successful in scheduling and y bursts fail to be scheduled. Out of these y bursts that have invoked rescheduling process, a fraction denoted by k has been successful in rescheduling. Now, $x + ky$ bursts are successful and total of $x + ky$ “RESERVE” control packets and ky “NOTIFY” control packets are sent. The fraction of increased signalling overhead is $\frac{ky}{x+ky}$. Consider a case where 90 of the total 100 bursts are successful while the remaining 10 fail to be scheduled. Further assuming that $k = 20\%$. The additional signalling overhead will therefore be $\frac{ky}{x+ky} = \frac{0.2*10}{90+0.2*10} = 0.0217 = 2.17\%$. It is clear that ODBR effectively does not incur significant signalling overhead.

3.4.2 Signalling Overhead for ABR

For ABR, although rescheduling would be invoked every time after a burst has been scheduled successfully, rescheduling is not feasible at all times. Consider a network where x bursts of the total bursts arrived have been successful in scheduling. Out of x bursts that have invoked rescheduling process, only a fraction, say, k' has been successful in rescheduling. The total number of control messages is therefore $x + k'x$, out of which x are of “RESERVE” type and $k'x$ are of “NOTIFY” type. The fraction of additional signalling overhead is $\frac{k'x}{x} = k'$. For a network state with $k' = 20\%$, the signalling overhead is 20%. However, as explained earlier, additional signalling overhead is not very significant when compared to the computational complexity gain achieved over LAUC-VF.

3.4.3 Signalling Overhead for BR-WFR

For BR-WFR, since rescheduling is done in the aggressive way (after each successful scheduling, the rescheduling process is invoked) rather than on-demand, the signalling overhead can be quantified in the same way as ABR.

3.5 Feasibility of Implementation

The implementation of rescheduling is feasible as explained below.

- Rescheduling of a burst is done before the burst actually arrives at a node. It does not affect any ongoing traffic.
- Wavelength reassignment does not change the time schedule of a burst on the outgoing link and on other links along the path. Only the next node needs to be notified about the change in the wavelength.

- FDL reassignment changes the time schedule. However, we restrict such a reassignment to a burst for which the outgoing link is the last hop. This means that only the next node (which is also the destination node for that burst) needs to be informed.

3.6 Performance Study

We evaluate the performance of ODBR and ABR in WDM OBS networks without any FDLs. Further, the performance of BR-WFR is evaluated in WDM OBS networks equipped with limited FDLs. The details of their respective performance study will be discussed in separate sections below. First, we discuss some details of our simulation model in Section 3.6.1 and the performance metrics used in Section 3.6.2.

3.6.1 Simulation Model

The simulation is performed using an event-driven simulator written using C programming language. We consider a random network of 32 nodes with 60 bidirectional links for the performance study of ODBR and ABR. A bidirectional link is realized by two unidirectional links in opposite directions. All the nodes are equally probable to be a destination node for a burst. The shortest path is used for routing a burst from the source to destination node. Those bursts that can not be scheduled are dropped. Bursts on the order of 10^5 are generated to obtain accurate results with 95% confidence interval of approximately 5% deviation from the reported mean value.

The burst durations are exponentially distributed with a mean (L) of $10 \mu s$. The control packet processing time (δ) is assumed to be $1 \mu s$. These values are chosen so that the ratio $\frac{L}{\delta}$ is greater than the ratio of the number of wavelengths used for data and control traffic. This will ensure that the control packet is processed and also transmitted before its data burst

is transmitted on a link even at full link utilization [14]. In the simulation, each link has 8 wavelengths for carrying data traffic and 1 wavelength for control traffic. The transmission capacity of each wavelength is approximately 10 *Gpbs*. The propagation delays between the node pairs ranges from 300 *ms* to 1000 *ms*.

We consider a multi-class case where two different classes of traffic namely class 1 traffic and class 2 traffic are generated. The average burst lengths of class 1 and class 2 traffic are denoted by L_1 and L_2 , respectively. For the performance study of ODBR and ABR, class 2 traffic is given a higher priority over class 1 traffic by assigning an extra offset time given by $3 * L_1$ as used in [28] as this could achieve 95% degree of isolation between the two different classes in wavelength reservation. In our experiments, we consider both class 1 and class 2 traffic having equal mean burst length, i.e. $L_2 = L_1 = L$. An initial offset time of δh is used for class 1 traffic while $\delta h + \text{extra offset}$ is used for class 2 traffic, where h is the number of hops along the route. Performance of class 1, class 2, and overall (combined) traffic for all the algorithms are studied.

3.6.2 Performance Metrics

We compare the performance of these algorithms with that of LAUC and LAUC-VF. The performance metrics used are *burst dropping probability*, *performance improvement*, and *effectiveness*. Burst dropping probability is defined as the ratio of the number of dropped bursts to the total number of bursts that arrive. If x bursts are dropped out of a total of y bursts, the dropping probability is given by $\frac{x}{y}$. Performance improvement indicates the percentage of improvement in dropping probability when compared to that of LAUC. The effectiveness is measured as the percentage of improvement in dropping probability when compared to that of LAUC as the upper bound and that of LAUC-VF as the lower bound. If the dropping probability of LAUC, ODBR (ABR or BR-WFR), and LAUC-VF is D_l , D_o , and D_v respec-

tively, the improvement of ODBR (ABR or BR-WFR) is $\frac{D_l - D_o}{D_l}$ and that of LAUC-VF is $\frac{D_l - D_v}{D_l}$, while the effectiveness of ODBR (ABR or BR-WFR) is $\frac{D_l - D_o}{D_l - D_v}$.

3.6.3 Performance study of ODBR and ABR

In this section, performance of the proposed ODBR and ABR algorithms is studied through simulation experiments. We compare the performance of the proposed algorithms under different traffic loading conditions. The burst arrival rate is measured as the number of bursts arrived per node per microsecond. The range for traffic load is chosen to be from 0.3 to 0.6 so that the burst dropping probability is below 15%.

Figures 3.8, 3.9, and 3.10, indicate that ODBR and ABR have better performance in terms of burst dropping probability than LAUC for overall, class 1, and class 2 traffic, respectively. The dropping probability increases with increasing traffic load as most of the wavelengths are heavily used at high traffic load, therefore, it is less probable for a burst to find an available wavelength. However, as shown in Figures 3.8 and 3.9, the performances of ODBR and ABR are always in between LAUC and LAUC-VF. Particularly, our proposed algorithms perform closer to LAUC-VF at low arrival rates than at high arrival rates. This is because, more voids are created at high arrival rates and our rescheduling algorithms consider only the last burst for rescheduling and do not utilize the voids in between the burst as in LAUC-VF.

Figure 3.10 shows that all algorithms have similar dropping performance for class 2 high priority traffic. This is because class 2 traffic have large initial offset time as compared to class 1 traffic, and it makes class 2 burst highly likely to be reserving wavelength at the far end on the time line. Therefore, a high priority burst is highly likely to be the last burst and hence not much improvement is achieved by LAUC-VF and also our algorithms over LAUC.

Figures 3.11, 3.12, and 3.13 show the percentage of improvement achieved with increasing

traffic load for overall traffic, class 1 traffic, and class 2 traffic, respectively. As for the overall traffic and class 1 traffic as shown in Figure 3.11 and 3.12, ABR and ODBR have about 32% and 35% performance improvement over LAUC, respectively at low traffic load as compared to 45% performance improvement achieved by LAUC-VF. The percentage improvement drops over increasing traffic load as more bursts have occupied the wavelengths and not many wavelengths are available for burst rescheduling. From Figure 3.13, we observe that class 2 high priority traffic for ABR and ODBR have shown about 5% and 9% performance improvement over LAUC while about 10% performance improvement has been achieved by LAUC-VF. The performance improvement achieved by ODBR, ABR, and LAUC-VF is low due to the large offset used by the high priority bursts as discussed above.

Figures 3.14, 3.15, and 3.16 depict the effectiveness of ODBR and ABR relative to LAUC and LAUC-VF under various traffic loads. As for the overall traffic performance and also the class 1 traffic performance as shown in Figure 3.14 and 3.15, effectiveness of about 75% and 70%, respectively, are achieved by ODBR and ABR, at low traffic load. The effectiveness of ODBR and ABR for class 2 traffic is plotted in Figure 3.16. In general, the effectiveness decreases with the increasing arrival rates. This is because more voids are created at high arrival rates and our rescheduling algorithms do not make use of these voids as in LAUC-VF.

Through simulation, we observed the number of “RESERVE” control packets that correspond to successful bursts and the number of “NOTIFY” control packets that correspond to successful rescheduling on each of the links. The results show that about 2% and 20% of signalling overhead is incurred by ODBR and ABR, respectively.

3.6.4 Performance study of BR-WFR

In this section, the performance of the proposed BR-WFR algorithm is studied through simulation with different traffic loading as well as different FDL size in a 32 nodes with 82

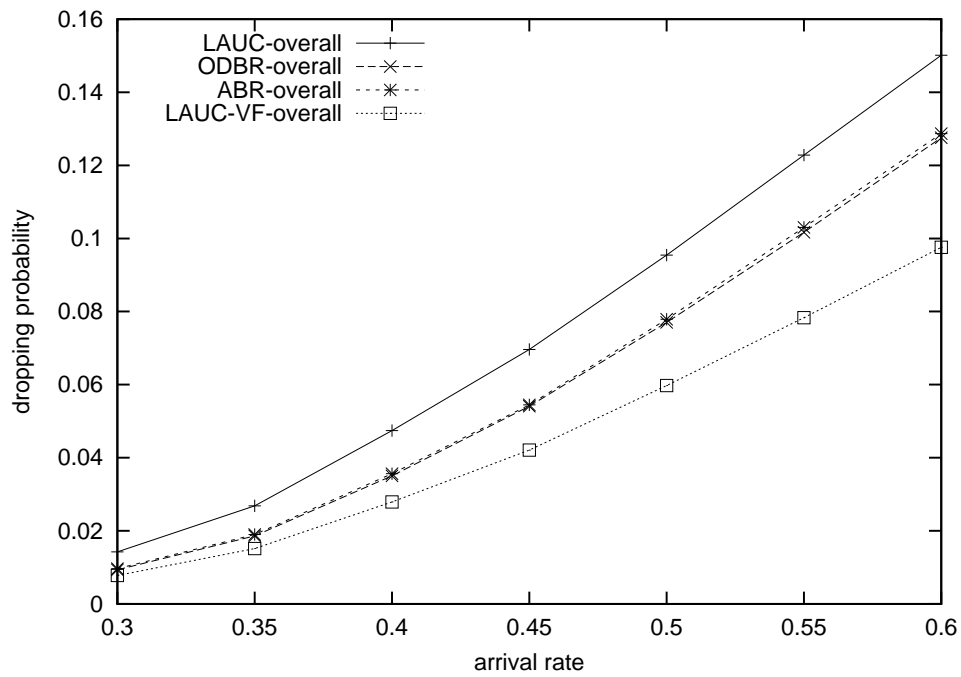


Figure 3.8: Performance of overall traffic for various algorithms under different traffic loading.

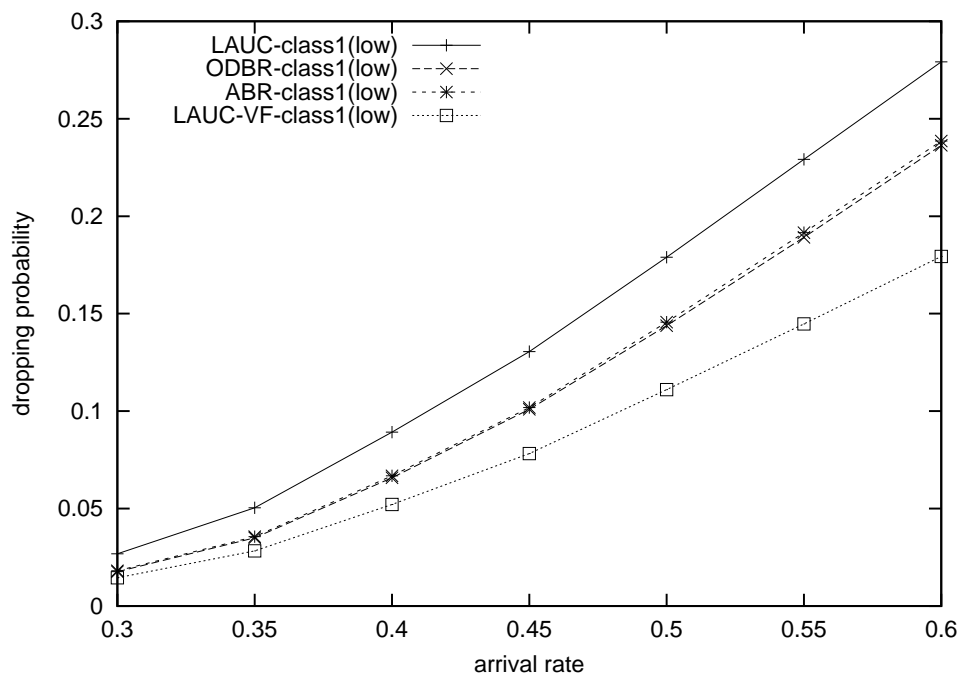


Figure 3.9: Performance of class 1 traffic for various algorithms under different traffic loading.

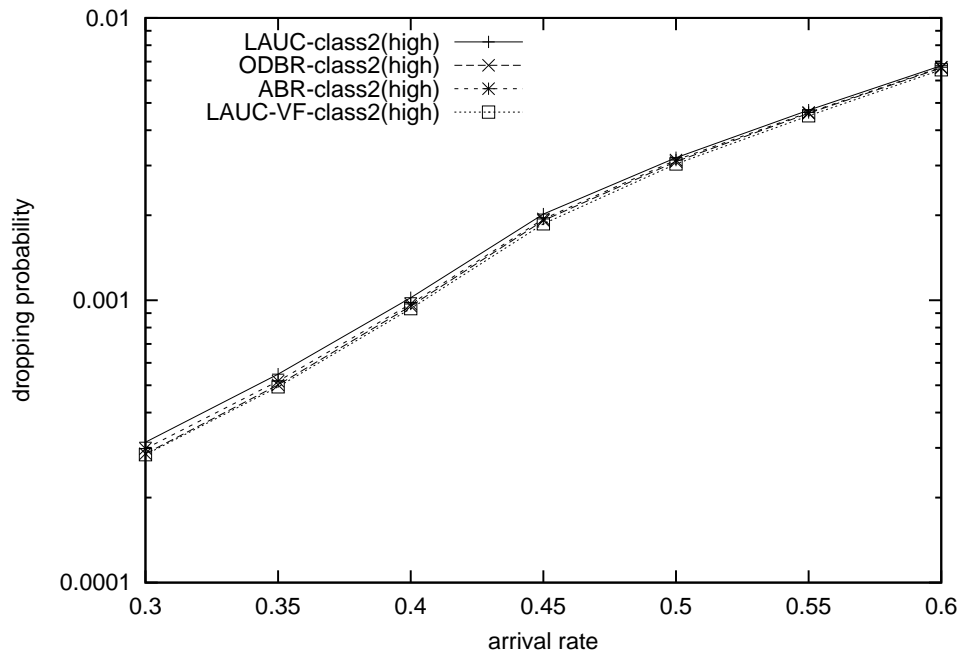


Figure 3.10: Performance of class 2 traffic for various algorithms under different traffic loading.

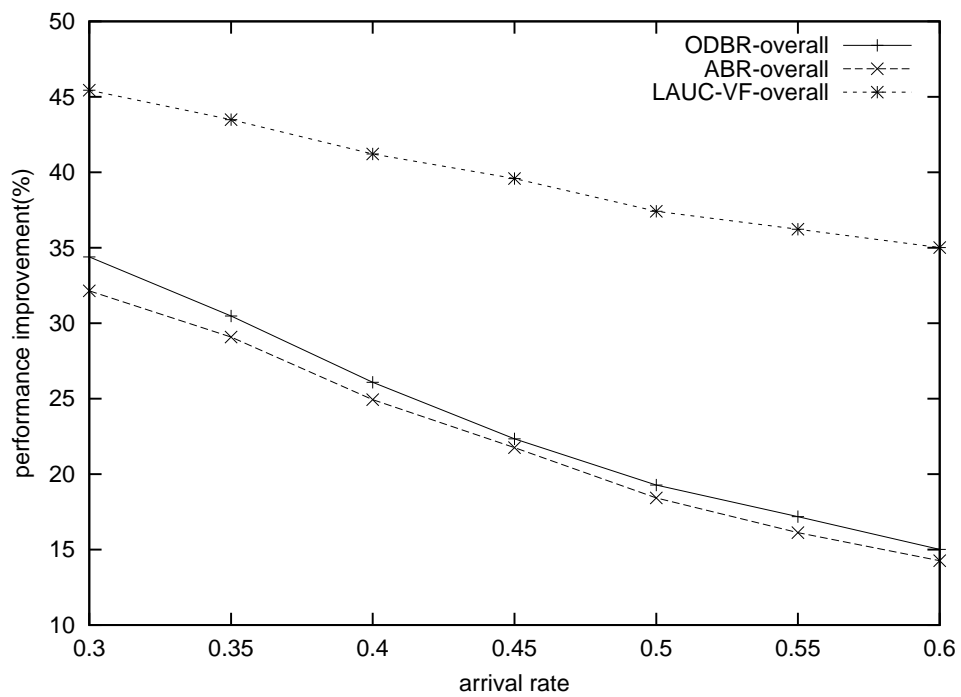


Figure 3.11: Performance improvement of overall traffic for various algorithms under different traffic loading.

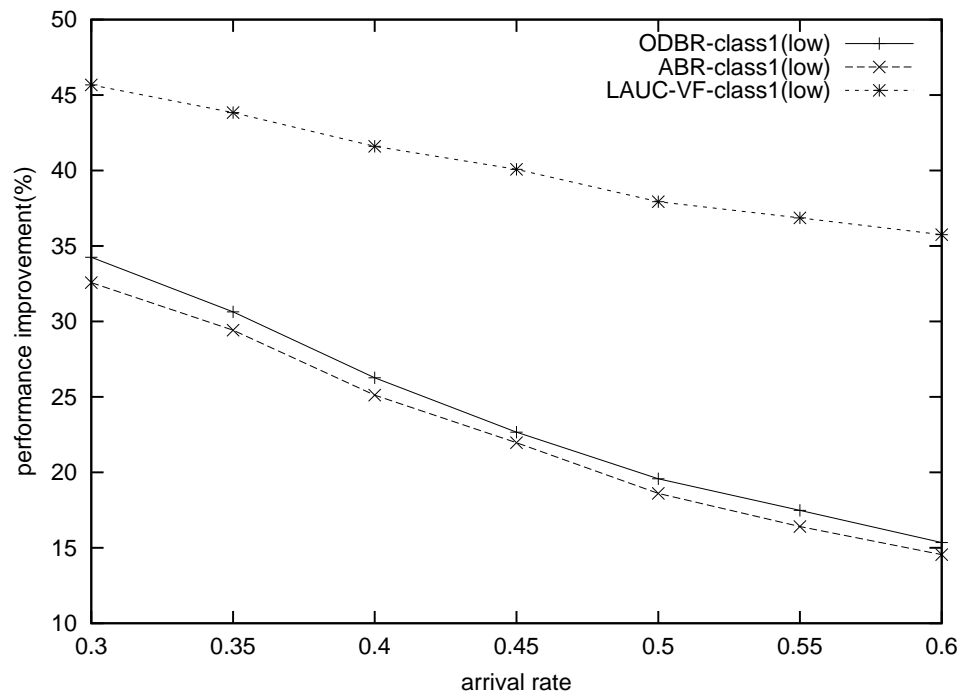


Figure 3.12: Performance improvement of class 1 traffic for various algorithms under different traffic loading.

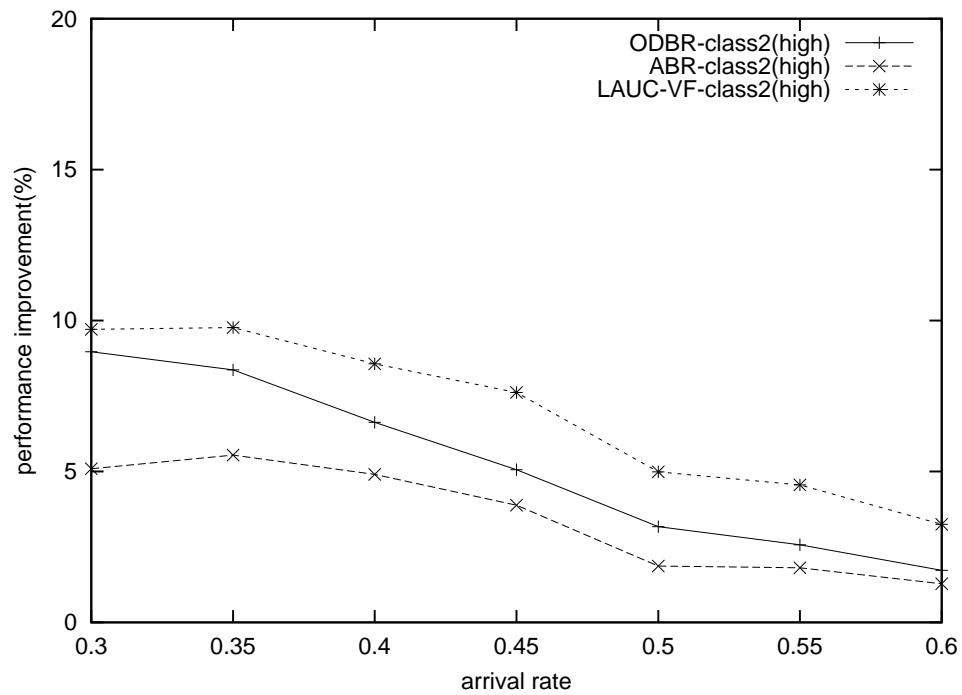


Figure 3.13: Performance improvement of class 2 traffic for various algorithms under different traffic loading.

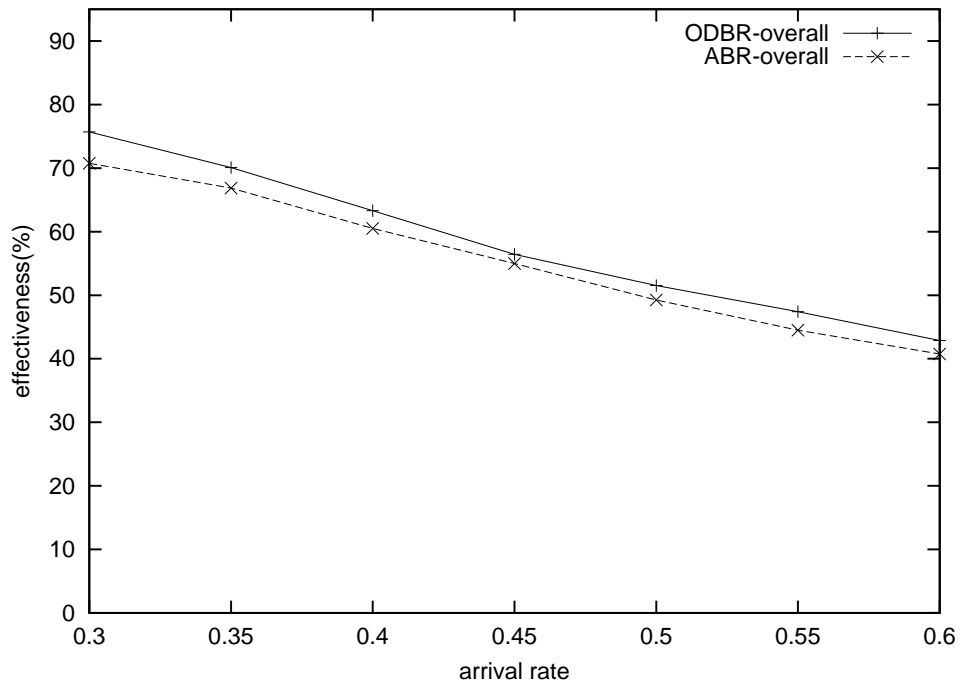


Figure 3.14: Effectiveness of overall traffic for ODBR and ABR under different traffic loading.

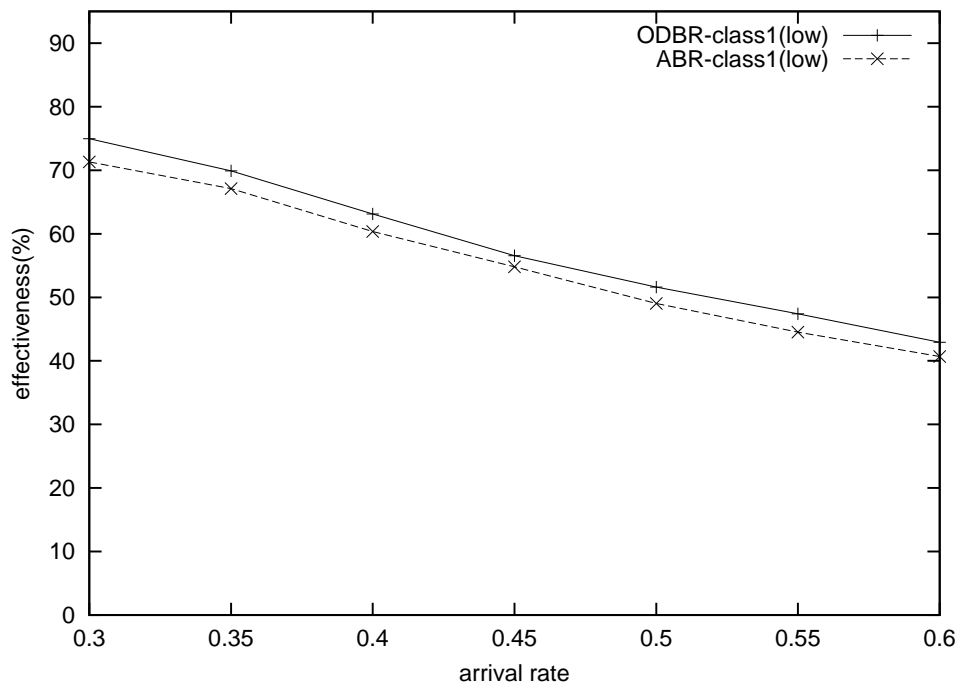


Figure 3.15: Effectiveness of class 1 traffic for ODBR and ABR under different traffic loading.

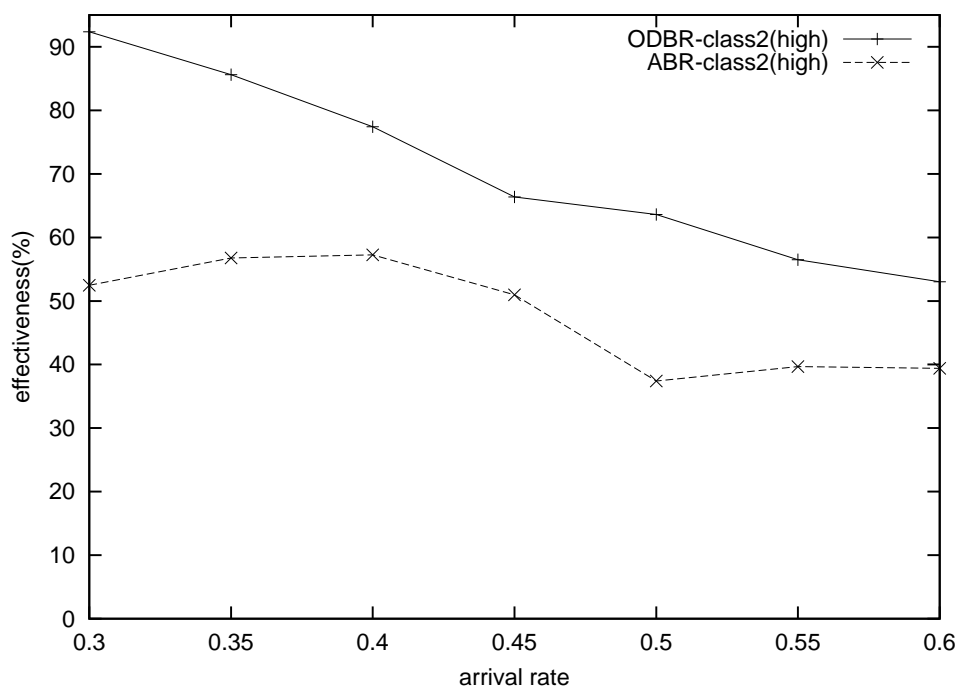


Figure 3.16: Effectiveness of class 2 traffic for ODBR and ABR under different traffic loading.

bidirectional links random network equipped with FDLs. Its performance is compared with that of LAUC and LAUC-VF in a random network equipped with FDLs. The FDL length is measured in units of μs . We consider two classes of traffic where higher priority class 2 requests are given extra offset time of $3 * L_1$ plus the maximum FDL length. Again, we consider $L_2 = L_1 = L = 10\mu s$ and $\delta = 1\mu s$. Other details of the simulation are the same as before. We also study the performance of the rescheduling algorithm BR-WR which differs from BR-WFR in that it carries out only wavelength reassignment but not FDL reassignment.

3.6.5 Effect of Traffic Loading

The dropping probabilities for the overall, class 1, and class 2 requests are shown in Figures 3.17, 3.18, and 3.19, respectively for different arrival rates (per node per μs) and an FDL buffer size of 10. As the arrival rate increases, the dropping probability for all the algorithms increases. LAUC experiences the highest dropping probability. BR-WFR and BR-WR per-

form better than LAUC and close to LAUC-VF at low traffic load. The performance of all the algorithms are similar for the class 2 high-priority requests as shown in Figure 3.19. This is due to the fact that a class 2 high-priority request has a higher offset time, which makes it more likely to be assigned a wavelength channel. Moreover, it is not affected by any void because of the long offset time used. Therefore, not much improvement can be obtained by using LAUC-VF or our proposed algorithms. Also, we observe that as expected class 2 requests perform far better than class 1 requests.

The performances of BR-WFR and BR-WR are better than that of LAUC for traffic loads ranging from low to high and are closer to that of LAUC-VF at low traffic load. Since in practice, networks usually operate at low traffic loads (smaller than 5% dropping probability), our algorithms are useful. This is shown in Figure 3.20, 3.21, and 3.22 where the performance improvements for the overall, class 1, and class 2 requests for BR-WFR, BR-WR and LAUC-VF over LAUC are presented. Figure 3.22 shows that the dropping probability for all the algorithms over LAUC have similar performance for the class 2 high-priority traffic. This is due to the large initial offset time used for these high-priority traffic, as explained in the previous paragraph.

BR-WFR performs better than BR-WR at low traffic load as shown in Figures 3.20, 3.21, and 3.22. With both wavelength reassignment and last-hop FDL reassignment, the chances of rescheduling bursts are higher and bursts are packed tighter. However, as the burst arrival rate increases, difference in the performance between the two algorithms diminishes. This is because, the possibility of finding eligible bursts for rescheduling decreases with increasing traffic load. Through simulation, we observed that the signalling overhead incurred by BR-WFR and BR-WR to be about 27% and 23%, respectively. The signalling packet is small in size as it needs to carry only the wavelength-change information to the next node. Therefore, the extra signalling packets would not take up much bandwidth on the signalling channel.

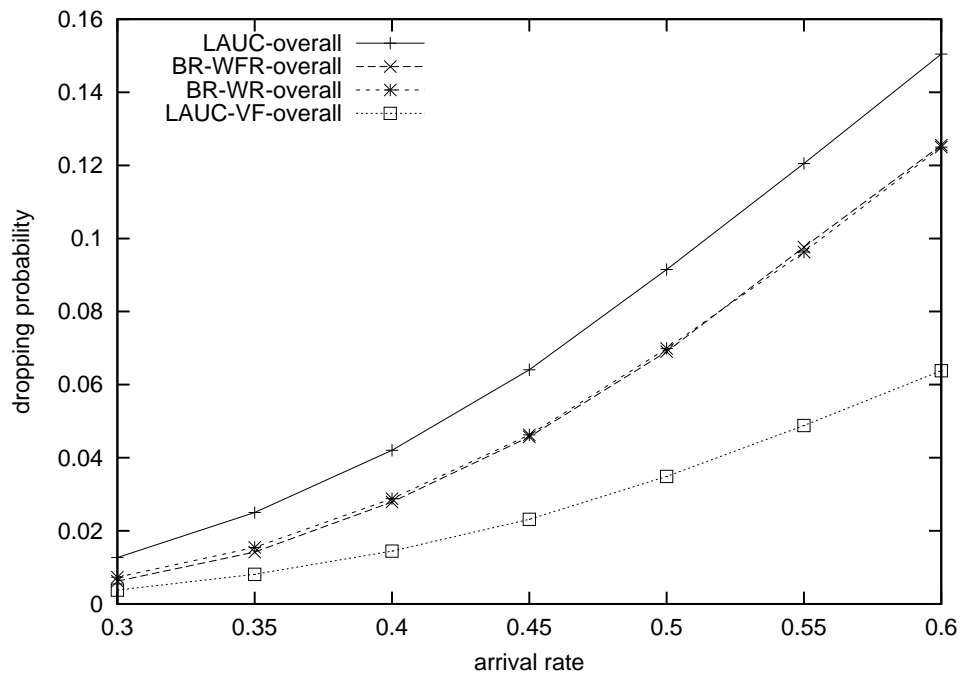


Figure 3.17: Performance of overall (class 1 and class 2) bursts for varying traffic load.

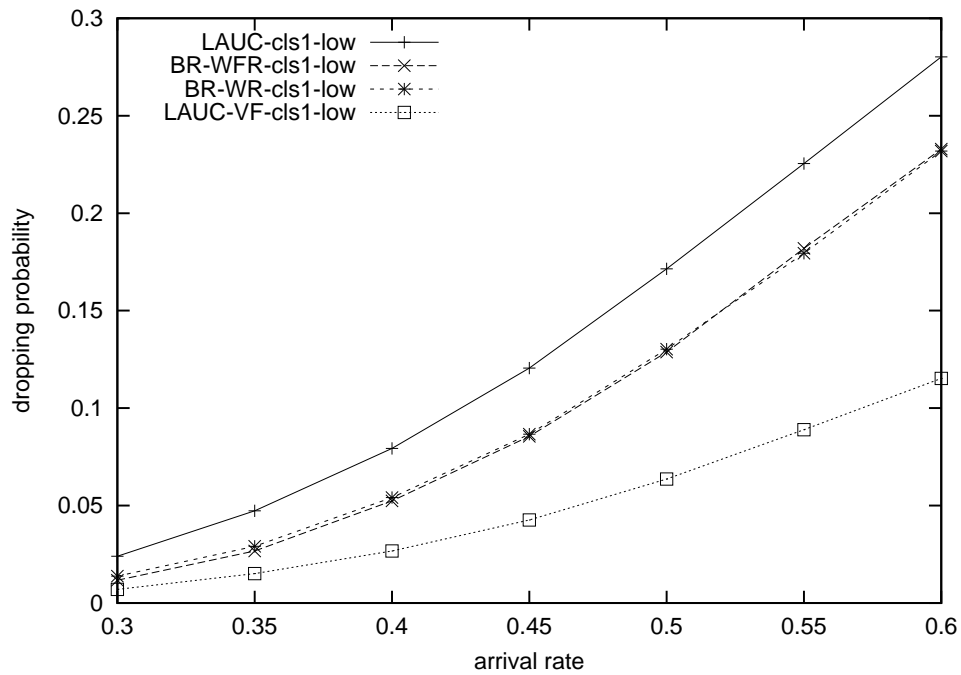


Figure 3.18: Performance of class 1 bursts for varying traffic load.

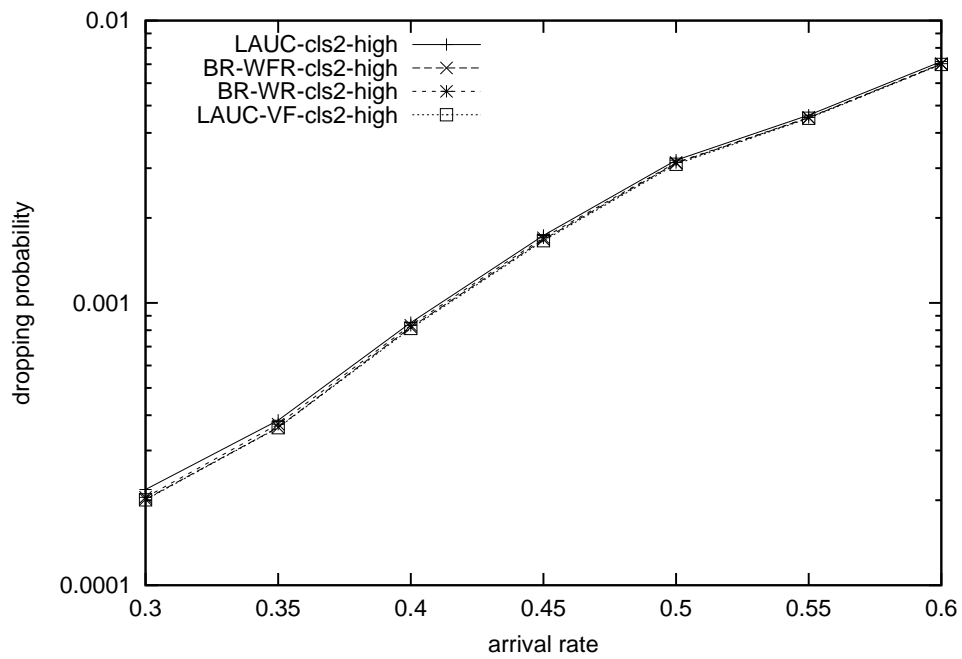


Figure 3.19: Performance of class 2 bursts with varying traffic load.

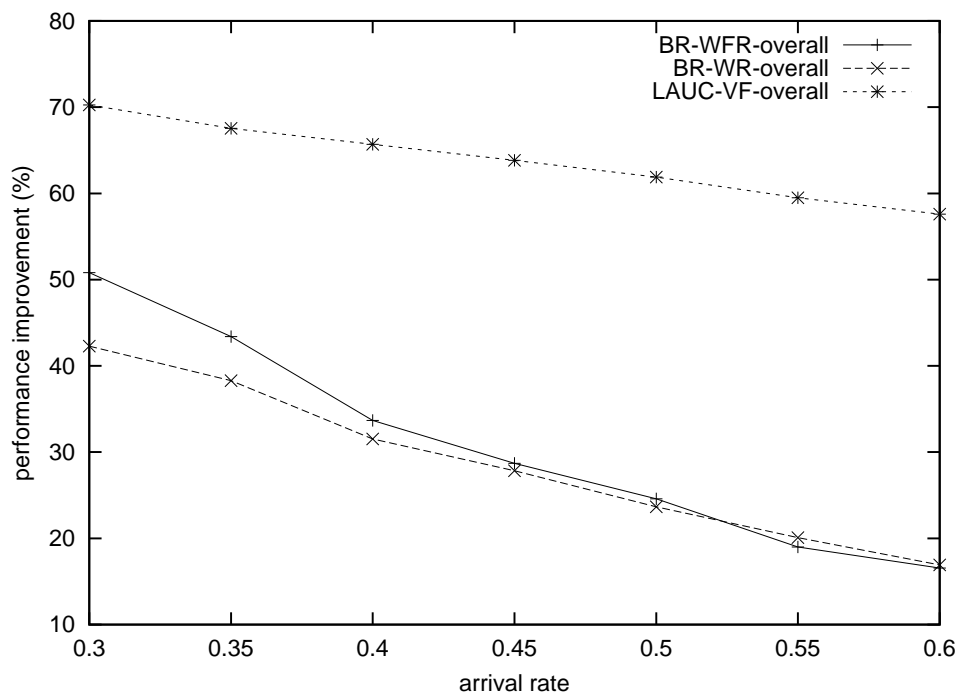


Figure 3.20: Performance improvement achieved by BR-WFR, BR-WR, and LAUC-VF over LAUC for overall bursts for varying traffic load.

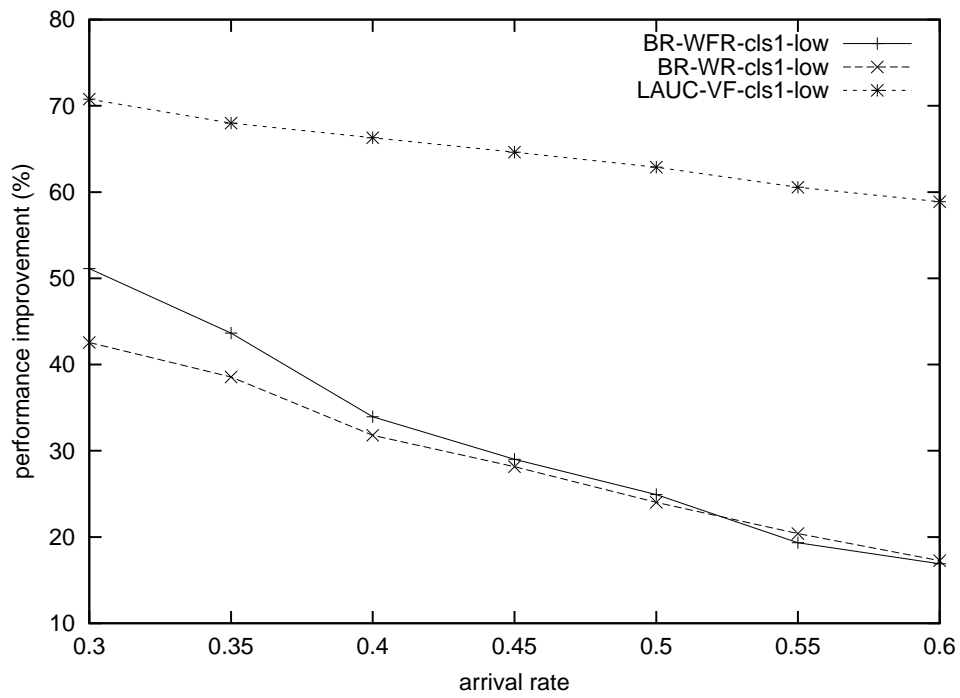


Figure 3.21: Performance improvement achieved by BR-WFR, BR-WR, and LAUC-VF over LAUC for class 1 bursts with varying traffic load.

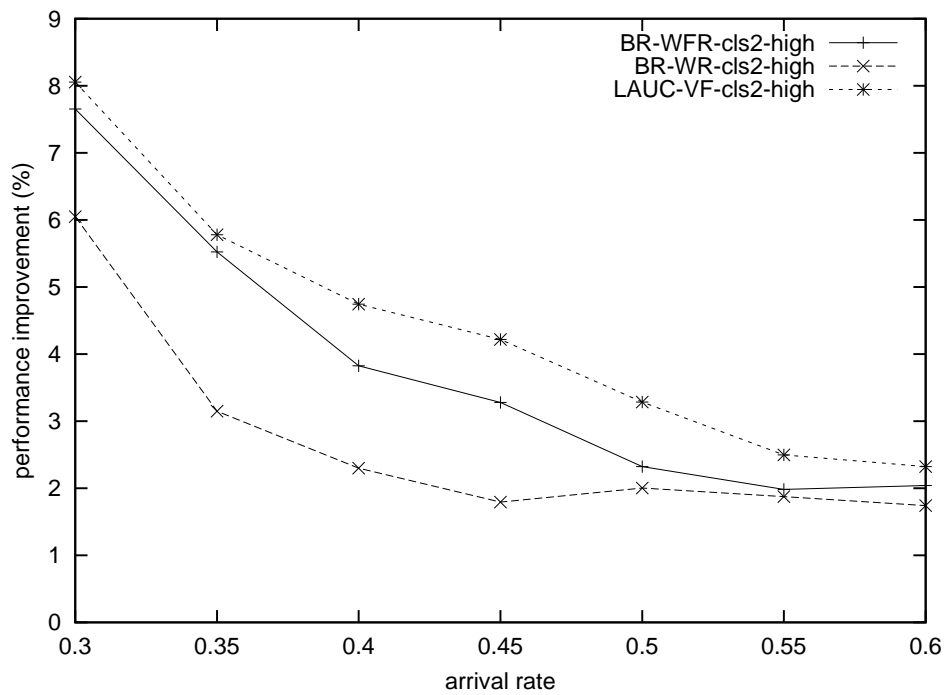


Figure 3.22: Performance improvement achieved by BR-WFR, BR-WR, and LAUC-VF over LAUC for class 2 bursts with varying traffic load.

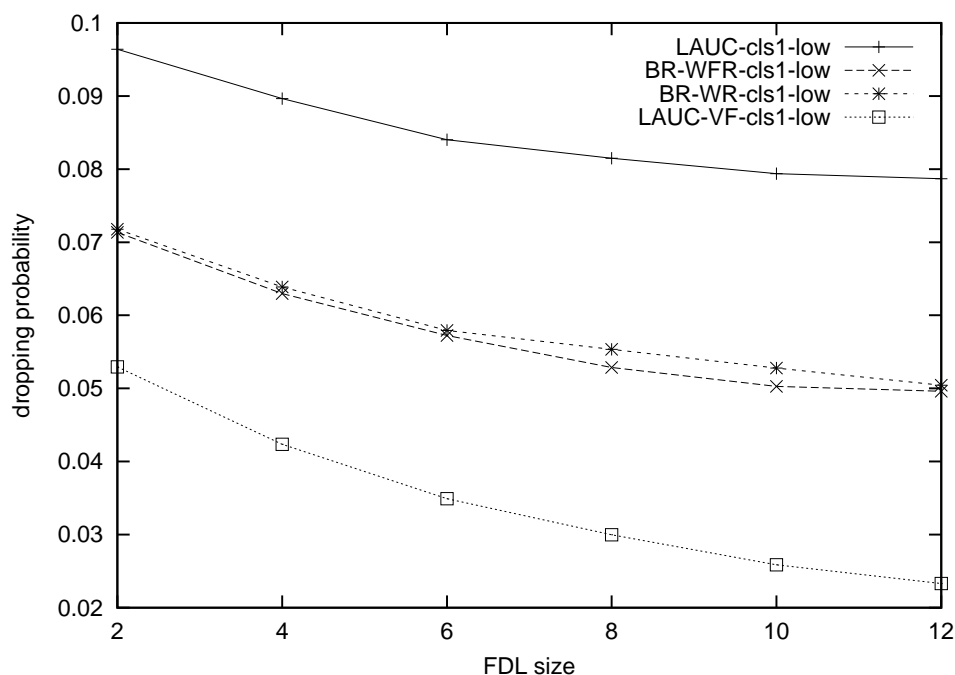


Figure 3.23: Performance of class 1 bursts for varying FDL size.

3.6.6 Effect of FDL Buffer size

We study the performance of BR-WFR and BR-WR for varying lengths (or size) of FDL buffer. The dropping probabilities of the algorithms with increasing FDL size for the class 1, and class 2 requests are shown in Figures 3.23, and 3.24, respectively, for an arrival rate of 0.4. In general, the dropping probability decreases with the increasing FDL size because with a long delay the possibility of finding a free wavelength increases. For the range of FDL size from 2 to 12, the signalling overhead incurred by BR-WFR and BR-WR algorithms fall in the range of 23% to 28% and 20% to 23% respectively.

3.7 Summary

In this chapter, we have proposed two rescheduling techniques, i.e., wavelength reassignment and last-hop FDL reassignment. We have proposed rescheduling algorithms for supporting

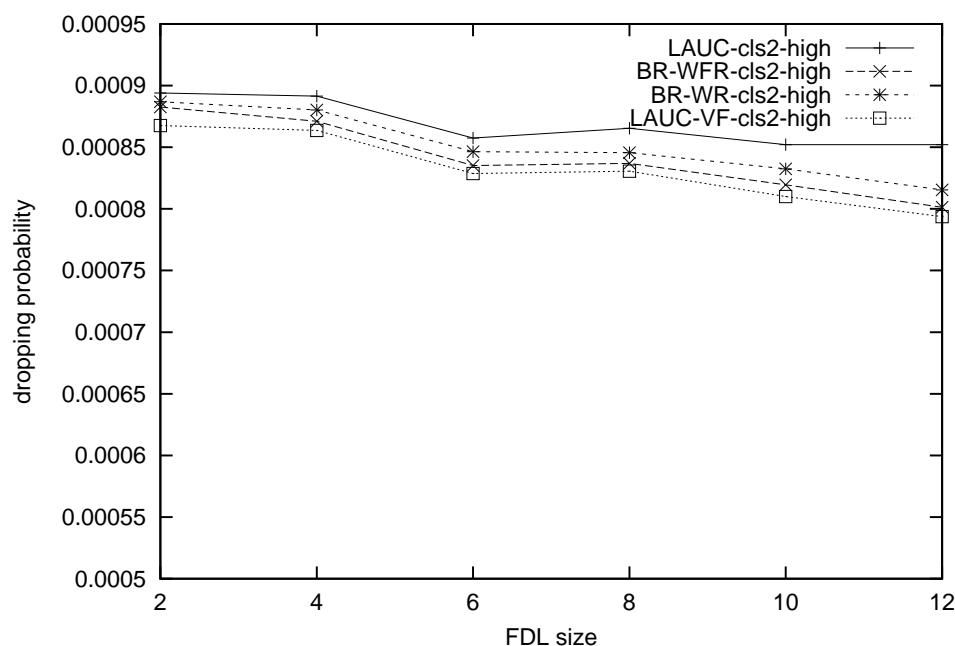


Figure 3.24: Performance of class 2 bursts for varying FDL size.

bursty traffic in optical networks with the objective of improving burst dropping performance keeping the computational complexity low. Since rescheduling takes place before a burst arrives, it does not disrupt any traffic. First, we developed rescheduling algorithms namely ODBR and ABR, make use of the wavelength reassignment technique. Since the rescheduling algorithm with wavelength reassignment changes only the wavelength, keeping the time unchanged, it does not pose any implementation problem. Further, we developed rescheduling algorithm, namely BR-WFR, which is suitable for networks equipped with limited FDLs. BR-WFR makes use of one or both of the two rescheduling techniques. As we restrict FDL reassignment to a burst for which the outgoing link is the last hop, only the destination node needs to be informed about the change in time schedule. Simulation results have shown that the proposed ODBR, ABR and BR-WFR algorithms perform significantly better than LAUC algorithm in terms of burst dropping probability. At the same time their performance is close to that of the existing complex LAUC-VF algorithm at low loads. The proposed algorithms, apart from supporting service differentiation efficiently in terms of burst dropping probab-

ity, have been shown to contribute a part to improve fairness. The proposed rescheduling algorithms have been shown to improve the performance of the low priority traffic significantly at low loads, thus alleviating the situation where low dropping of the high priority traffic is realized at the expense of the high dropping of low priority traffic in an extra offset time based multi-class environment. Further, the signalling overhead incurred by the proposed algorithms has been studied and observed to be less significant when compared to the computational complexity gain achieved over LAUC-VF.

Chapter 4

Offset Management for Fairness Improvement

In this chapter, we consider the hop length based *fairness* issue in OBS networks. Fairness here refers to, for all ingress-egress node pairs in a network, a burst having equal likelihood to get through independent of the hop length (or hop count) involved. In WDM-based OBS networks, bursts that traverse longer paths are more likely to be dropped compared to bursts that traverse shorter paths, resulting in a fairness problem. This is because longer-hop bursts have a higher probability of not finding a free wavelength on a link (e.g., higher chances of encountering a bottleneck link). The fairness problem is more pronounced in OBS networks because of the lack of optical buffers at the core nodes. The fairness problem can also be caused by other factors such as variations in burst lengths as described in [75]. Our work focuses on the fairness problem due to hop lengths.

The fairness problem occurs in classless as well as in multi-class environments. As described in [28], different offset time values are assigned to different traffic classes in order to achieve service differentiation in pJET. A larger offset time is used by a higher priority burst so that

its control packet can reserve resources further in advance, which increases its chances of getting the required bandwidth resources. However, within a class, bursts with longer hop lengths but without sufficient offset times still experience higher dropping probabilities than those with shorter hop lengths. A variation of JET called JET-FA has been proposed in [12] to achieve fairness. The key idea is to assign an extra offset time proportional to the hop count, e.g., $5hL$ or $10hL$, where L is the mean burst duration. This extra offset time does not take into account the traffic load, network state and topology. For example, with $5hL$, $L = 20 \mu s$, and $h = 5$, an extra offset time of $0.5 ms$ is needed. This long initial offset time increases the end-to-end delay and requires large buffers at the ingress nodes to queue the bursts. Shorter-hop bursts also tend to be over-penalized as they have much shorter offset times. Also, this method is only applicable to classless traffic and cannot be directly extended to multi-class traffic with varying priorities.

In this chapter, we develop an efficient fairness method called *link scheduling state based offset selection (LSOS)*. This method is suitable for networks with no FDLs at the core nodes and is suitable for a classless as well as multi-class environment. The objective of LSOS is to manage the offset times and choose offset times based on the link states for bursts with different hop lengths such that they perform almost equally. The state of a link is defined by the link scheduling probabilities for different possible offset times. The link scheduling probabilities are computed at the core nodes and collected periodically from the core nodes via link state advertisements by the edge nodes which then determine the offset times needed for bursts traversing different hop lengths. Apart from ensuring fairness without over-penalizing the shorter-hop bursts by using online link states, LSOS makes explicit routing with sufficiently minimum offset time possible. Further, by limiting the offset range for various classes, it ensures that the initial offset time and hence the queueing delay at the edge nodes is at an acceptable level. The effectiveness of the proposed method is evaluated through simulation experiments.

The rest of this chapter is organized as follows. An overview of the proposed LSOS method is presented in Section 4.1, followed by the details of the method in Section 4.2. Section 4.3 is devoted to the performance study. Concluding remarks are made in Section 4.4.

4.1 Overview of LSOS

A burst with a longer hop path has a higher dropping probability, resulting in the fairness problem. The selection of appropriate initial offset times in order to ensure fairness is non trivial. LSOS chooses the amount of initial offset times needed at an ingress node for bursts with different hop lengths based on online link states. The collection of the link states is done for a limited time and is used for a subsequent time period much longer than that of the collection period. The basic assumption is that link states normally do not change abruptly. Indeed, it has been accepted in other traffic engineering methods, e.g., [76, 77, 78], that the aggregate traffic on links normally changes very slowly as observed in [79, 80]. Also studies [81] have shown that traffic in IP networks often exhibits long-range dependence, with the implication that congested periods can be quite long. Therefore, a good representation of the network state may be obtainable through limited time measurements and this state remains valid for a relatively longer time period. This implies the processing (link state collection) overheads at the core nodes and the overheads incurred in the exchange of link states will not be significant. Link states are measured at a core node by observing bursts at its outgoing link for various possible offset time values. The maximum value of the initial offset time used is directly related to the maximum delay bound and is subject to the decision of the network operator. Link states are advertised to the edge nodes periodically so that they can adapt to the dynamically fluctuating network state over time.

By using the link state information, an ingress node selects the initial offset times for bursts

traversing different hop lengths. For instance, for a 2-hop path, the initial offset time needed is decided by considering the link states on the first and second links. The state of a link is defined by the link scheduling probabilities for different offset times. The link scheduling probability for a link at some offset τ is the probability of finding a free wavelength on the link to schedule a burst at time $t + \tau$, where t is the current time. The path scheduling probability is defined by multiplying the scheduling probabilities of the constituent links at appropriate offset times. If T is the offset time used at the first link then the reduced offset time of $T - \delta$ is used at the second link where δ accounts for the control processing time. By considering a few possible values for T , the path scheduling probabilities of a 2-hop path are obtained. The offset time T which gives the path scheduling probability for this 2-hop path closest to that of a 1-hop path is selected to be the initial offset time. By extending this method to offset time selection for bursts with different hop counts, LSOS makes sure that the burst dropping performances for all bursts regardless of their hop lengths are almost equal.

As an example, consider a 2-hop path 1-2-3 as shown in Figure 4.1. Assume that the control packet processing time at a node is 2 time units. The link scheduling probabilities for links 1-2 and 2-3 for different offset time values are shown in Figure 4.1. The scheduling probability of the 1-hop path at the minimum required initial offset time is chosen as the reference probability value. In this example, this value is 0.9681 which corresponds to an offset of 2 time units. In Table 4.1 the scheduling probabilities of the 2-hop path 1-2-3 are given for various offset values. The minimum offset required is 4 units at node 1 which implies an offset value of 2 units at node 2. By multiplying the corresponding scheduling probability at the links 1-2 and 2-3 at these offset values, we obtain the path scheduling probability of $0.9805 \times 0.9581 = 0.9394$. Since this value is much lower than that of the 1-hop path, 2-hop bursts are more likely to be dropped than 1-hop bursts leading to the fairness problem. From Table 4.1, it can be observed that an initial offset time value of 8 units gives the path scheduling probability which is the closest to the reference probability value than any other

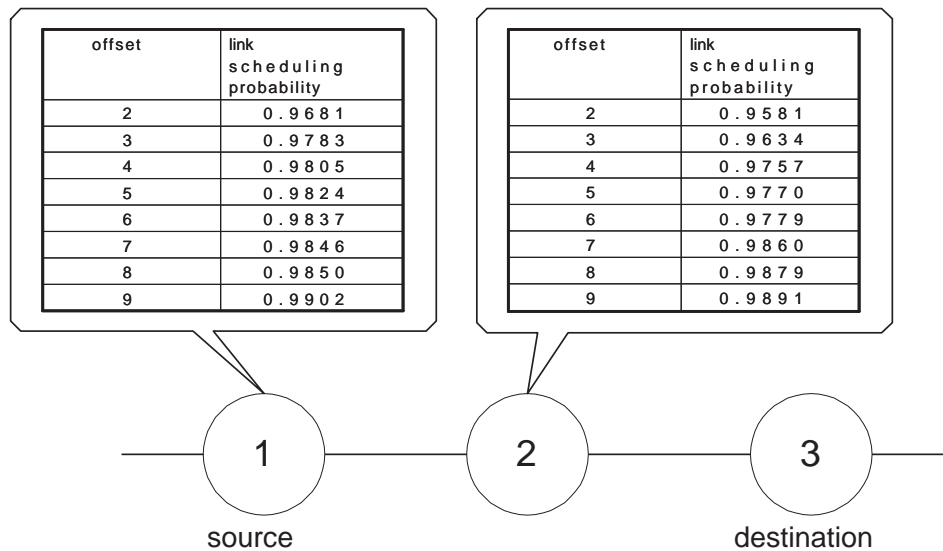


Figure 4.1: Link states on a 2-hop path.

offset time values. Hence, LSOS chooses 8 units as the initial offset time for bursts traversing the path 1-2-3.

As shown in the example, link states on all h links of a h -hop path are used in calculating the path scheduling probability. Therefore, every core node needs to maintain the scheduling states for all its outgoing links. All link states are advertised to the edge nodes periodically. A possible way of reducing the signalling overhead caused by link states advertisements is by considering $k < h$ links, so that not all link states on all links are used. The link states collected on the k links are extrapolated to the remaining links. A possible way is to use the scheduling probability of the bottleneck link (which is the one involving the minimum value of scheduling probability) among the first k links for the remaining links.

Before presenting the details of LSOS in the next section, the following first list the attractive features of the LSOS method.

- Explicit routing is done with sufficient offset times for ingress and egress node pairs with different hop lengths.

Table 4.1: 2-hop path scheduling probability for different offset time values.

Offset at node 1	Offset at node 2	Path scheduling probability
4	2	0.9394
5	3	0.9464
6	4	0.9598
7	5	0.9620
8	6	0.9632
9	7	0.9763

- Fairness is achieved with an offset time within a predefined range.
- Traffic loads and network connectivity are taken into consideration since online link state information is used.
- Signalling overhead is small as link state collection is done for a short time period only and the offset times computed are used for a sufficiently longer time period thereafter.
- By using the scheduling state information on a few links only, the need for global state information can be avoided.

4.2 LSOS for Intra-class Fairness

4.2.1 Preliminaries

With LSOS, the progression in time can be viewed in terms of cycles where each cycle consists of two phases, a *link-probing phase* followed by a *new-offset phase*. For a cycle with time period T_p , the link-probing phase lasts for a limited time t_c and the new-offset phase lasts for a much

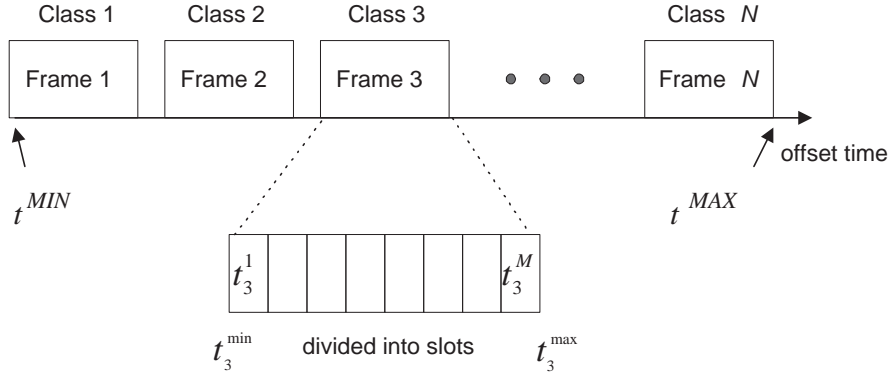


Figure 4.2: Division of offset time into frames for different priority classes of traffic.

longer time $T_p - t_c$. In the link probing phase, the outgoing link is probed repeatedly to check the availability of wavelengths at different possible offset times. At the end of this phase, a new set of *link state tables* are generated at the core nodes. The link state tables are collected by the edge nodes via link state advertisements. Based on these tables *hop-offset tables* are generated locally at every edge node. A hop-offset table indicates the initial offset time needed for a burst traversing a h -hop path during the new-offset phase until the next set of offset times is computed.

To support differentiated services, bursts are categorized into N classes denoted by C_m , $1 \leq m \leq N$. The range of offset values, F , with maximum and minimum values given by $F = [t^{MIN}, t^{MAX}]$, is divided into N frames as shown in Figure 4.2. One frame is assigned to one burst class with frame F_m assigned to class C_m . This is to distinguish the priorities of bursts through offset times with higher offset times assigned to higher priority bursts. Within class C_m , the minimum and maximum offset times are $F_m = [t_m^{min}, t_m^{max}]$. To ensure isolation of classes, a gap is set between any two frames. Each offset time frame is further divided into M slots. A burst belonging to class C_m will have an offset time that coincides with a slot in frame F_m . The slot s of frame F_m for class C_m is denoted by F_m^s . We use the following notations and definitions to describe LSOS.

Notations

N : Number of burst classes;

h : Number of hops in a path;

C_m : Burst class type m ;

F_m : The offset time frame for bursts of class C_m ;

F_m^s : Slot s in frame F_m ;

M : Number of slots in a frame;

t^{MAX} : The maximum offset time value for offset time frame F ;

t^{MIN} : The minimum offset time value for offset time frame F ;

t_m^{MAX} : The maximum offset time value for offset time frame F_m ;

t_m^{MIN} : The minimum offset time value for offset time frame F_m ;

l_q : A link indexed q ;

R^L : A route/path traversing a set L of links.

Definitions

$$C = \{C_m | m = 1, 2, \dots, N\};$$

$$F = \{F_m | m = 1, 2, \dots, N\};$$

$$F = [t^{MIN}, t^{MAX}];$$

$$F_m = \{F_m^s | s = 1, 2, \dots, M\};$$

$$F_m = [t_m^{min}, t_m^{max}];$$

$$L = \{l_q | q = 1, 2, \dots, h\};$$

$$R^L = R^{l_1, l_2, \dots, l_h}.$$

The procedure for assigning offset time frames to different priority classes is given in Table 4.2.

Table 4.2: Offset Time Assignment to Different Priority Classes

- input: N classes of bursts with class type denoted by $C = \{C_m | m = 1, 2, \dots, N\}$.
- Step 1: Offset time F with a range of offset values given by $[t^{MIN}, t^{MAX}]$ is divided into N frames, where $F = \{F_m | m = 1, 2, \dots, N\}$. All N frames have respective minimum and maximum values given by $F_m = [t_m^{min}, t_m^{max}]$ where $t_{m-1}^{max} < t_m^{min}$, $1 < m \leq N$.
- Step 2: Assign $F_m = [t_m^{min}, t_m^{max}]$ to C_m , $m = 1, 2, \dots, N$, where $t_1^{min} = t^{MIN}$ and $t_N^{max} = t^{MAX}$.
- Step 3: Each frame is divided into M slots such that slots in F_m are given by $F_m^s = t_m^{min} + sz$, $s = 1, 2, \dots, M$, where M and z are predefined.

4.2.2 Computation of Link Scheduling Probabilities

Without loss of generality, assume (slot size) $z = \delta$ where δ is the control packet processing time. At a node, whenever a burst of class C_m arrives, the outgoing link is observed to determine if there is a free wavelength at time $t = F_m^1, t = F_m^2 \dots t = F_m^M$. At the end of a predefined link-probing phase, a link state table is generated locally at each node as shown in Figure 4.3. The generation of the link state table is based on the link scheduling probabilities. The link scheduling probability for link l_q denoted by $P_m^s(l_q)$ is the probability that the burst is successful in finding a free wavelength at time $t = F_m^s, s = 1, 2, \dots, M$. Whenever a burst with length b_l arrives, a virtual burst with length b_l is used to calculate the link scheduling probabilities at different offset times. Thus $P_m^s(l_q)$ is calculated as $P_m^s(l_q) = x_m^s(l_q)/x_m(l_q)$, where $x_m(l_q)$ is the total number of virtual bursts of class C_m used and $x_m^s(l_q)$ is the number of class C_m bursts that can possibly be successful in finding a free wavelength with offset time F_m^s on link l_q . The larger the offset time, t , the earlier the reservation is made, therefore it is highly likely that $P_m^1(l_q) \leq P_m^2(l_q) \dots \leq P_m^M(l_q)$.

The pseudocode for generating a link state table for link l_q for class C_m is given in Table 4.3. For simplicity, the subscript m is omitted.

4.2.3 Offset Selection

An ingress node makes use of the link states to generate a hop-offset table which indicates the offset times needed for bursts traversing a h -hop route denoted by R^L . Let $P(R^L)$ be the path scheduling probability of a burst traversing route R^L . In order to achieve fairness, the initial offset time, $T = F_m^j$ is chosen such that $P^j(R^L) = P^j(l_1) \times P^{j-1}(l_2) \times \dots \times P^{j-h+1}(l_h)$ has the closest value to a reference probability. Here, this reference probability is $P^1(R^{l_1})$ which corresponds to the success probability for a 1-hop burst with the same ingress node.

Table 4.3: Computation of Link Scheduling Probabilities

Step 1: Initialization {beginning of link-probing phase}

```

 $x(l_q) \leftarrow 0$ 

for  $s = 1, 2, \dots, M$  do
     $x^s(l_q) \leftarrow 0$ 
endfor

```

Step 2: Upon arrival of a burst with length b_l {link-probing phase}

```

 $x(l_q) \leftarrow x(l_q) + 1$ 

for  $s = 1, 2, \dots, M$  do

    if a burst with length  $b_l$  can be virtually scheduled on link  $l_q$  at
     $t = F^s$ 

    then  $x^s(l_q) \leftarrow x^s(l_q) + 1$ 

    endif

endfor

```

Step 3: Calculation of link scheduling probabilities {end of link-probing phase}

```

for  $s = 1, 2, \dots, M$  do

     $P^s(l_q) \leftarrow x^s(l_q)/x(l_q)$ 

endfor

```

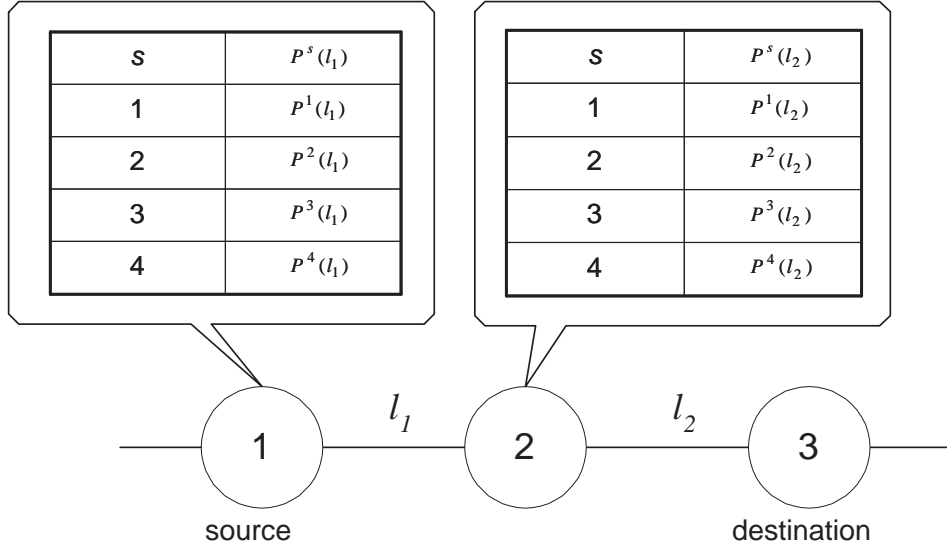


Figure 4.3: Illustration of link state tables generated at nodes.

In the following discussion, the subscript m is omitted for simplicity. Consider a 2-hop path as shown in Figure 4.3 with link l_1 and link l_2 . In order to calculate the best offset time needed for a burst originating from node 1 and destined for node 3, let offset $t_{off} = F^s$. For several possible values of s , the success probability for a h -hop burst between the source and destination nodes is calculated as $P^s(R^L) = P^s(l_1) \times P^{s-1}(l_2) \times \dots P^{s-h+1}(l_h)$. The value of s is chosen such that $t_{off} \geq h\delta$, which is the minimum initial offset time needed. The ultimate goal is to find the best candidate j of s such that $P^j(R^L)$ is the closest to $P^1(R^{l_1})$. Therefore, $P^s(R^{l_1, l_2}) = P^s(l_1) \times P^{s-1}(l_2)$. Since link states on all links along a route are used, this method of offset selection is called all-LSOS (A-LSOS). The pseudocode for determining the offset value t_{off} for a h -hop path for class C_m is given in Table 4.4. P_{ref} is the reference probability that corresponds to a 1-hop path.

To reduce link state advertisement overheads, link states on $k < h$ links can be collected and extrapolated to the remaining links along the route. A possibility is to choose the bottleneck link among these k links as the reference link and use the state of this bottleneck link on all the other links along the route. For a given offset s , the bottleneck link is the link x ,

Table 4.4: Offset Time Selection

Step 1: Let s' be the minimum value such that the offset $F^{s'} \geq h\delta$

Step 2: $s \leftarrow s'$

$$diff \leftarrow \infty$$

$$P_{succ} \leftarrow 1$$

Step 3: for $i = 1$ to h do

$$P_{succ} \leftarrow P_{succ} \times P^{s-i+1}(l_i)$$

endfor

Step 4: if $|P_{succ} - P_{ref}| < diff$ then

$$diff \leftarrow |P_{succ} - P_{ref}|$$

$$t_{off} \leftarrow F^s$$

endif

Step 5: if $s < M$ then

$$s \leftarrow s + 1$$

goto Step 3

endif

Step 6: Choose t_{off} for the h -hop path

Table 4.5: A-LSOS and 1-LSOS Path Scheduling Probability of path 1-2-3.

	Path scheduling probability	
s	A-LSOS	1-LSOS
2	$P^2(R^{l_1, l_2}) = P^2(l_1) \times P^1(l_2)$	$P^2(R^{l_1, l_2}) = P^2(l_1) \times P^1(l_1)$
3	$P^3(R^{l_1, l_2}) = P^3(l_1) \times P^2(l_2)$	$P^3(R^{l_1, l_2}) = P^3(l_1) \times P^2(l_1)$
4	$P^4(R^{l_1, l_2}) = P^4(l_1) \times P^3(l_2)$	$P^4(R^{l_1, l_2}) = P^4(l_1) \times P^3(l_1)$

$x = 1, 2, \dots, k$, having the smallest link scheduling probability. To further reduce the amount of link states needed, k can be set to 1 so that the first link becomes the reference link. In this case, as only the link state on of the first link is used, the scheme is referred to as 1-LSOS. For example, Table 4.5 shows three different values for $P^s(R^{l_1, l_2})$ for A-LSOS and 1-LSOS respectively as a result of three different values of s . The value of s for which $P^s(R^{l_1, l_2})$ is the closest to $P^1(R^{l_1})$ (i.e. $P^1(l_1)$) is chosen. The associated offset time is chosen to be the initial offset time needed for a 2-hop burst originating from node 1.

4.3 Performance Study

The performance of LSOS has been studied through simulation on the 14-node NSFNET topology shown in Figure 4.4. Each fiber is assumed to have 8 wavelengths for data traffic and 1 wavelength for control traffic. The transmission capacity of each wavelength is approximately 10 Gpbs. Burst arrivals to the network are assumed to follow a Poisson process [28, 29, 82] and burst lengths are exponentially distributed with an average of 16 μs . The control packet processing time has been assumed to be 2 μs . Explicit routing with fixed shortest path has been used for routing a burst from its source node to its destination node.

The performance of LSOS has been studied for the four scenarios of classless traffic, multi-class traffic, identical traffic and non-identical traffic. In the identical traffic case, the traffic load is the same for all source and destination node pairs. In the non-identical traffic case, different node pairs have different traffic arrival rates (*burst*/ μsec). The effect of different link-probing phase periods on the performance of LSOS with classless and multi-class identical traffic has also been investigated.

The performance metrics used are the *burst dropping probability* and *standard deviation in burst dropping probability*. Burst dropping probability is calculated by taking the ratio of the total number of dropped bursts to the total number of bursts having the same hop length. The standard deviation in dropping probabilities encountered by bursts with different hop lengths has been used as a measure of fairness.

The burst scheduling algorithm used is the LAUC-VF algorithm. However, it is worth noting that LSOS fairness method functions independent of the scheduling algorithm used. Since with LSOS, different offset time ranges are assigned to different classes of traffic, bursts of various classes arriving at an OBS node which attempt to reserve the wavelength on the outgoing link will have different offset times. Therefore, voids are created on the wavelengths as time progresses as in the case of pJET, burst rescheduling algorithms proposed in Chapter 3 are likely to provide similar performance improvement over LAUC when used with LSOS.

4.3.1 Performance of LSOS in a Classless Traffic Environment

The performance of LSOS in a classless traffic environment with both identical and non-identical traffic demands is now presented with reference to the performance of the JET protocol. For the purpose of link states collection, a frame with 20 slots and slot size of $2\ \mu\text{s}$ (same as a control packet processing time) has been used. The possible offset times corresponding to this choice of frame size range from $2\ \mu\text{s}$ to $40\ \mu\text{s}$. As shown in Figure

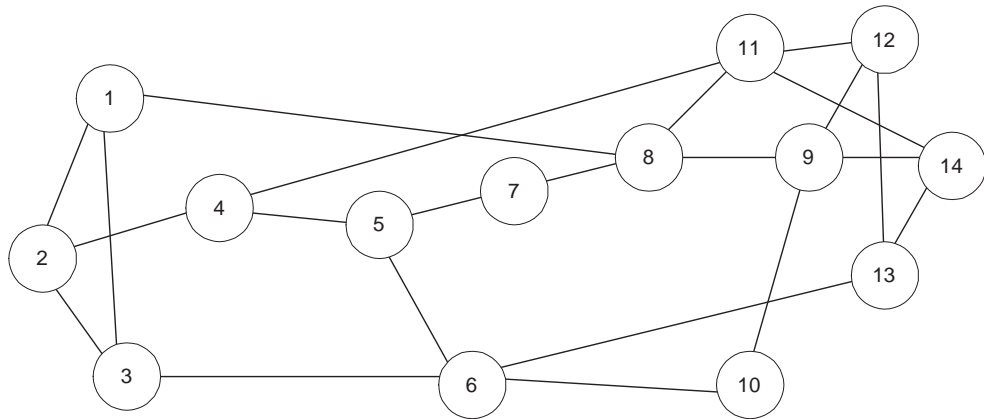


Figure 4.4: 14-node NSFNET.

4.5, with a burst arrival rate of 0.2, the dropping probability of JET increases with the hop length. The proposed A-LSOS and 1-LSOS methods on the other hand show better dropping performance than JET with increasing hop length. Figure 4.6 shows that with non-identical traffic where the arrival rates range from 0.15 to 0.3 among the nodes pairs, LSOS still achieves better dropping performance than the JET protocol with increasing hop length.

Table 4.6 shows the standard deviation in dropping probabilities for A-LSOS, 1-LSOS, and JET for the identical and non-identical traffic demands. It can be observed that both A-LSOS and 1-LSOS have lower values of standard deviation than JET with both identical and non-identical traffic. This shows the effectiveness of LSOS. Also, A-LSOS has better performance in terms of standard deviation in dropping probabilities than 1-LSOS. This is to be expected since the states of all links have used to provide a more accurate representation of the network state. As far as fairness performance in terms of standard deviation is concerned, A-LSOS always performs better than 1-LSOS. Since A-LSOS and 1-LSOS use different sets of offset times, in terms of the dropping probability 1-LSOS may perform better than A-LSOS.

Table 4.7 shows the mean offset time (in μs) needed for different hop lengths for A-LSOS and 1-LSOS with identical and non-identical traffic demands. The offset times needed for JET and JET-FA with different hop lengths are also presented in the same table for comparison.

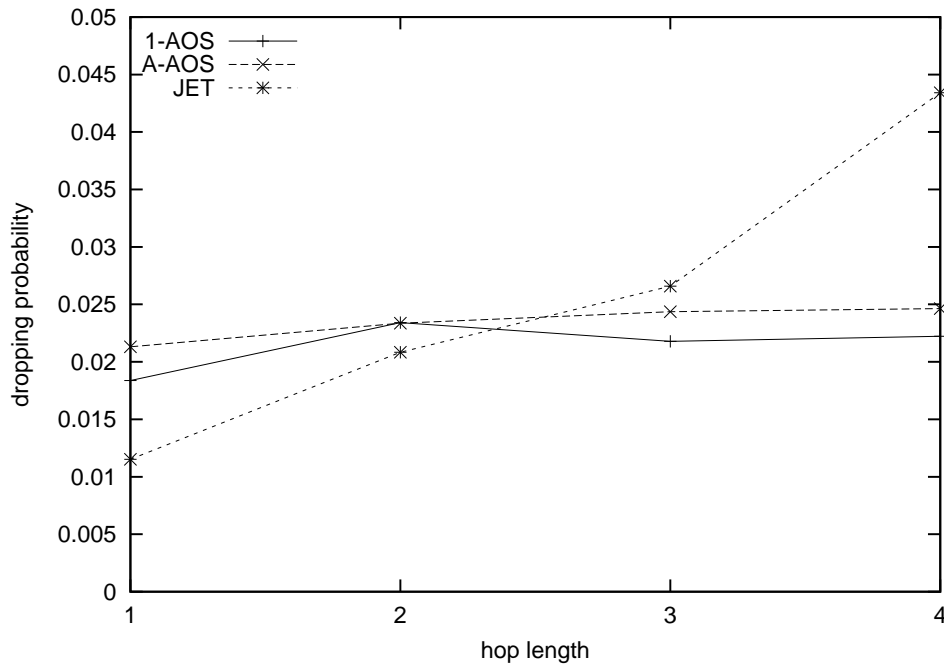


Figure 4.5: Dropping performance vs. hop length for classless traffic with identical traffic demand.

JET-FA achieves fairness by applying an extra offset time of $5h * 16$ to all the bursts in addition to the initial offset time. Since LSOS ensures fairness while JET does not, A-LSOS and 1-LSOS require higher offset times than JET. However, it is shown that A-LSOS and 1-LSOS require much lower offset times than JET-FA. It can also be observed that A-LSOS has a higher offset time value than 1-LSOS because link states of all links along a route are used. Hence, the chances of taking into account the state of a bottleneck link are higher.

Table 4.6: Standard deviation in burst dropping probabilities with different hop lengths for classless environment.

Traffic	A-LSOS	1-LSOS	JET
Identical	1.51E-03	2.16E-03	1.34E-02
Non-identical	2.31E-03	3.45E-03	1.64E-02

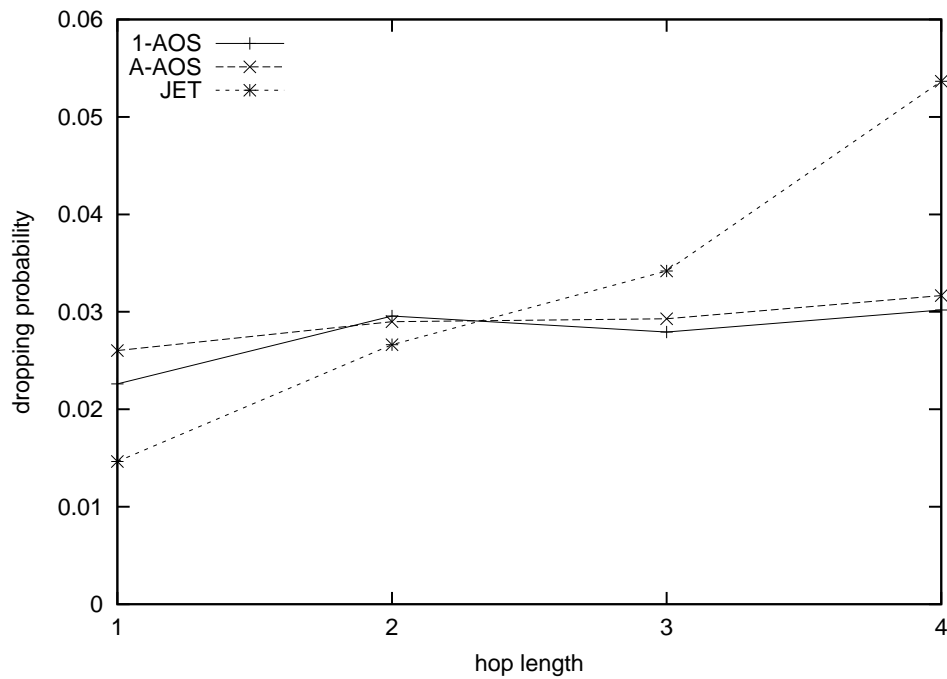


Figure 4.6: Dropping performance vs. hop length for classless traffic with non-identical traffic demand.

Table 4.7: Mean offset time (in μs) needed for A-LSOS, 1-LSOS, JET, and JET-FA for classless environment.

	Mean offset time					
	Identical traffic		Non-identical traf- fic		Any traffic	
Hop length	A-LSOS	1-LSOS	A-LSOS	1-LSOS	JET	JET-FA
1	2.00	2.00	2.00	2.00	2.00	82.00
2	15.34	10.40	14.80	10.76	4.00	164.00
3	22.00	16.80	20.66	16.60	6.00	246.00
4	28.00	20.86	25.00	16.80	8.00	328.00

4.3.2 Performance of LSOS in a Multi-class Environment

To study the performance of LSOS in a multi-class environment, two different traffic classes, class 1 traffic and class 2 traffic, with the same mean burst length value are considered. Class 2 traffic has higher priority compared to class 1 traffic. Priority is given by using a higher offset time value. Each offset time frame (with respect to each class of traffic) has 20 slots and the size of a slot is $2 \mu s$. The possible offset times for class 1 traffic range from $2 \mu s$ to $40 \mu s$. class 2 traffic has the possible offset times ranging from $60 \mu s$ to $98 \mu s$. Results of A-LSOS and 1-LSOS are compared with those for pJET. Under pJET, class 1 traffic has initial offset time T_{init} of δh and class 2 traffic has initial offset time of $\delta h + 3 * 16$.

As shown in Figures 4.7 and 4.8, with an arrival rate of 0.24, the dropping probability of pJET increases with increasing hop length. This shows the unfairness problem using pJET in a multi-class environment. The dropping performances of A-LSOS and 1-LSOS are better than that of pJET for both class 1 and class 2 traffic. This shows that fairness is achievable by the proposed LSOS algorithm for different classes of traffic in a multi-class environment. Particularly, fairness achieved by LSOS is better in class 2 high priority bursts than class 1 low priority bursts. This is because higher priority bursts are assigned higher offset times and this increases their chances of finding free wavelengths. Specifically, higher priority bursts traversing longer hop paths (which have higher probability of encountering a bottleneck link) benefit from being assigned with a larger offset time. In a non-identical traffic environment with the traffic load ranging from 0.15 to 0.3, the proposed LSOS algorithms show similar performance as shown in Figures 4.9 and 4.10.

Table 4.8 shows the standard deviation in dropping probabilities for A-LSOS, 1-LSOS, and pJET under identical traffic with the arrival rate of 0.24 and non-identical traffic with an arrival rate ranging from 0.15 to 0.3. Both A-LSOS and 1-LSOS have lower standard deviation values than pJET, which shows that improved fairness is achieved. Particularly A-LSOS

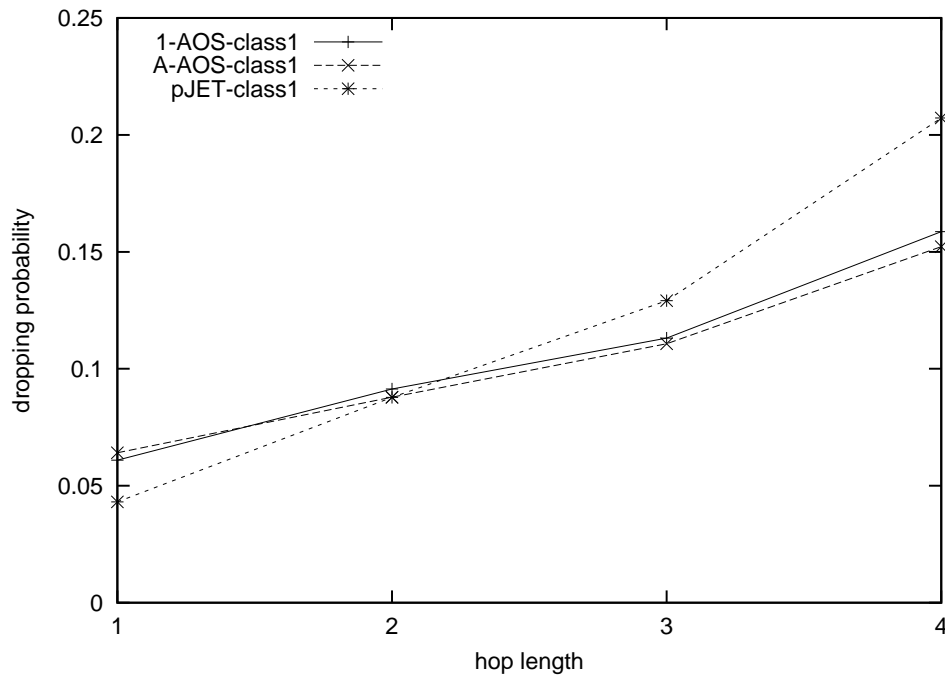


Figure 4.7: Dropping performance vs. hop length for class 1 traffic with identical traffic demand.

performs better than 1-LSOS in terms of fairness as all link states along a route have been used for the computation of the offset time needed.

Table 4.9 shows the mean offset time needed for different hop lengths for both A-LSOS and 1-LSOS under identical and non-identical traffic for class 1 and class 2 bursts. The offset time needed by pJET for different hop lengths for both classes have been stated for comparison. A-LSOS and 1-LSOS have higher mean offset times than pJET since pJET does not provide fairness. Again, A-LSOS has a higher offset time value than 1-LSOS as A-LSOS has a higher chance of involving a bottleneck link. Offset time values for class 2 traffic are similar for both A-LSOS and 1-LSOS since class 2 high-priority bursts are assigned an offset frame with larger offset values and the chances of finding a free wavelength for longer hop bursts become high with a small extra offset time itself. Compared to low-priority bursts, with higher offset time, high priority bursts outperform when finding free wavelengths.

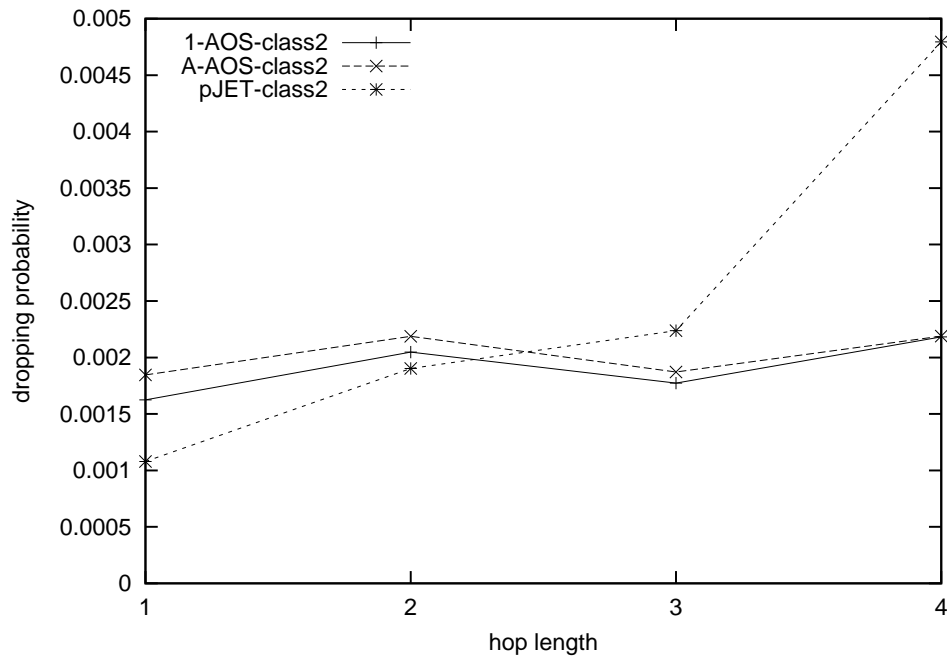


Figure 4.8: Dropping performance vs. hop length for class 2 traffic with identical traffic demand.

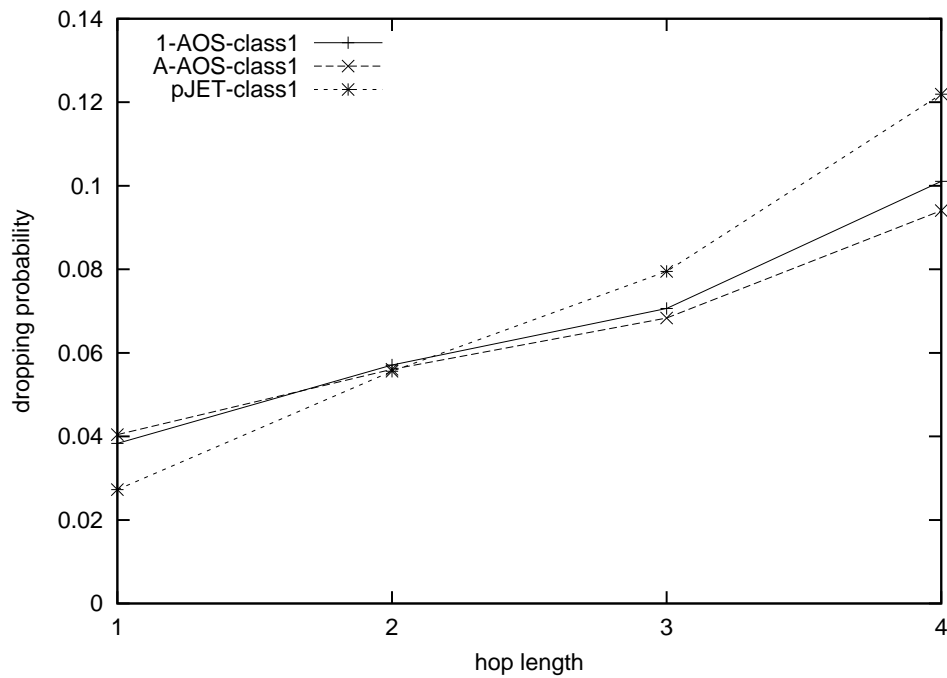


Figure 4.9: Dropping performance vs. hop length for class 1 traffic with non-identical traffic demand.

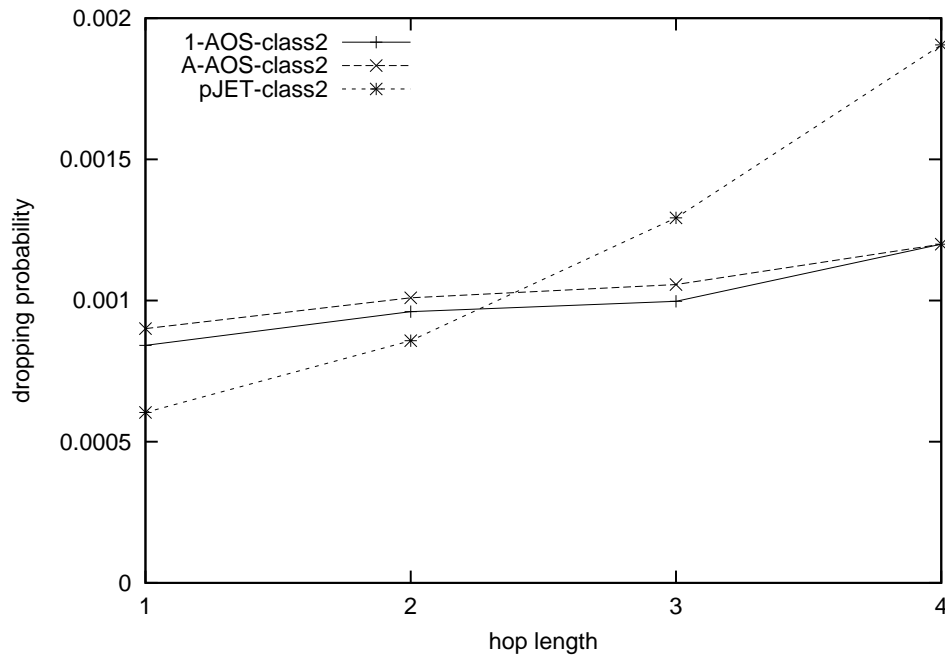


Figure 4.10: Dropping performance vs. hop length for class 2 traffic with non-identical traffic demand.

Table 4.8: Standard deviation in burst dropping probabilities of traffic with different hop lengths.

Traffic	Class	A-LSOS	1-LSOS	pJET
Identical	1	3.75E-02	4.11E-02	6.98E-02
	2	1.90E-04	2.54E-04	1.60E-03
Non-identical	1	2.27E-02	2.64E-02	4.01E-02
	2	1.24E-04	1.49E-04	5.70E-04

Table 4.9: The mean offset time (in μs) needed for A-LSOS, 1-LSOS, and pJET in multi-class traffic with different hop lengths.

		Mean offset time				
		Identical traffic		Non-identical traffic		Any traffic
Class	Hop length	A-LSOS	1-LSOS	A-LSOS	1-LSOS	pJET
1	1	2.00	2.00	2.00	2.00	2.00
	2	14.32	10.00	14.34	10.34	4.00
	3	20.66	15.00	22.00	16.00	6.00
	4	27.00	19.34	28.00	20.34	8.00
2	1	60.00	60.00	60.00	60.00	50.00
	2	66.00	66.00	66.00	66.00	52.00
	3	70.00	70.00	70.00	70.00	54.00
	4	74.00	74.00	74.00	74.00	56.00

4.3.3 Effect of the Link-probing Phase Period on the Performance of LSOS

The effect of the link-probing phase period on the performance of LSOS is studied by comparing the performance of 1-LSOS and A-LSOS with increasing link-probing period in a classless traffic environment followed by a multi-class traffic environment. Identical traffic demands with the same set of parameters as before are considered. From Figure 4.11, we observe that with increasing link-probing period, 1-LSOS and A-LSOS both perform better in terms of fairness as indicated by the decreasing standard deviation in dropping probabilities. The standard deviation for 1-LSOS and A-LSOS tend to stabilize at higher values of the link-probing period. This is because once sufficient information is obtained from the network state, the benefits of extra probing become less obvious when the network state does not change substantially. This indicates that long link-probing period is not necessarily needed.

Figure 4.12 shows the performance of class 1 low priority traffic of 1-LSOS and A-LSOS with increasing link-probing period. It is shown that the fairness performance of 1-LSOS and A-LSOS both improve as the link-probing period increases. Similar trend is observed for class 2 high priority traffic as shown in Figure 4.13. This shows the effectiveness of LSOS in a multi-class traffic environment. Again, the standard deviation for 1-LSOS and A-LSOS tend to stabilize at higher link-probing period due to the reason mentioned above.

4.4 Summary

In this chapter, an efficient link scheduling state based fairness improvement method to select the initial offset times needed for bursts with different hop lengths using the link scheduling probabilities has been proposed. As online link states are used, this method inherently captures the dynamic traffic loading and topological connectivity of the network. The per-

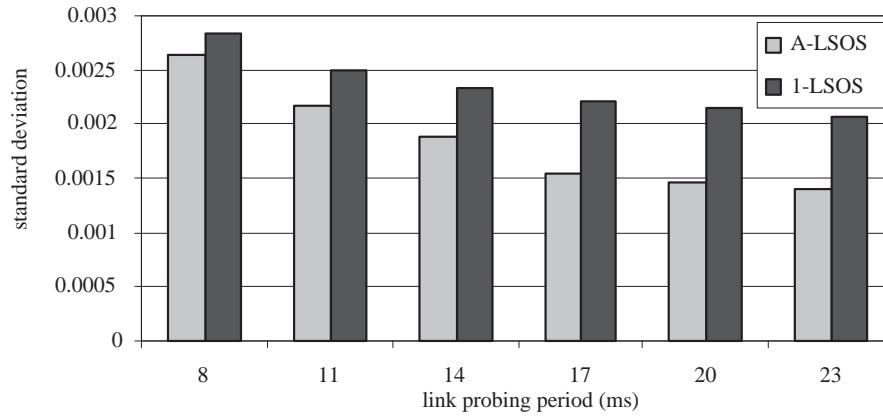


Figure 4.11: Standard deviation vs. link probing period for classless traffic.

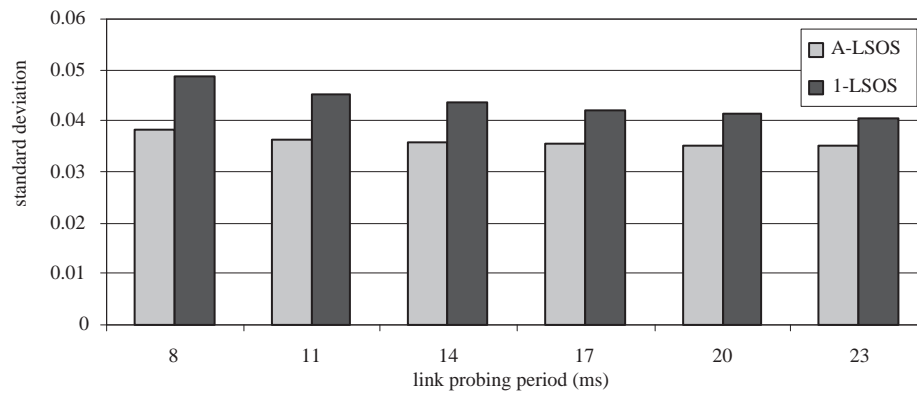


Figure 4.12: Standard deviation vs. link probing period for class 1 traffic.

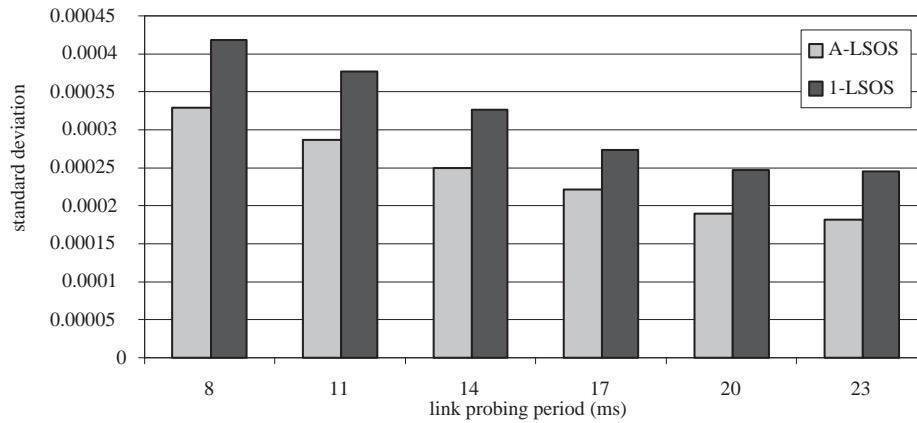


Figure 4.13: Standard deviation vs. link probing period for class 2 traffic.

formance of the proposed k -link based LSOS (k -LSOS) has been studied with A-LSOS and 1-LSOS where A-LSOS uses link states on all links along a route while 1-LSOS uses only the link state on the first link. The effectiveness of the proposed method has been demonstrated for classless and multi-class environments with identical and non-identical traffic demands. Improvement in fairness is achieved with a predefined acceptable range of offset times.

Chapter 5

Edge-to-Edge Proportional QoS Provisioning

In this chapter, we develop a feedback-based offset time selection (FOTS) method to provide proportional service differentiation in terms of burst dropping probability. Proportional QoS has recently received much attention as it allows a network operator to make adjustments on performance spacing between classes quantitatively in order to facilitate pricing. With proportional QoS, the QoS metric is adjusted to be proportional to the differentiation factor. For example, if d_i is the burst dropping metric and s_i is the differentiation factor for class i , the following proportional equation will hold for every one of N classes.

$$\frac{d_i}{d_j} = \frac{s_i}{s_j} \quad i, j = 1, \dots, N \quad (5.1)$$

Existing proportional schemes such as “intentional dropping” developed in [33] provide per-hop QoS between different classes of traffic in a proportional manner. As bursts are intentionally dropped regardless of the availability of required resources, the scheme results in poor bandwidth utilization. Further, guaranteeing per-hop proportional loss does not guarantee edge-to-edge proportional loss [34, 42]. This is because the edge-to-edge burst loss probability

involves multiplication of constituent loss probabilities on individual hops, thereby causing the edge-to-edge burst dropping ratio to deviate from the per-hop burst dropping ratio.

FOTS, by making use of link states collected by probe packets, dynamically adjusts the offset times needed to achieve the predefined proportional QoS among different classes of traffic for various ingress-egress node pairs. As the offset time selection is done for the node pairs, fairness among node pairs with various hop lengths in terms of achieving the proportional QoS is maintained. The effectiveness of the proposed method is evaluated through simulation experiments. In this research work, we consider burst dropping probability as the QoS metric. Also, the ratio of the differentiation factor is referred to as the predefined proportional loss ratio.

The rest of this chapter is organized as follows. In Section 5.1, an analysis on providing proportional QoS with offset time for a single link model is presented. This analysis provides some insights on how these two concepts (of proportional QoS and offset time) work together, providing the basis for the proposed FOTS method. This is followed by the description of the proposed FOTS method in Section 5.2. Section 5.3 presents the performance study of the proposed FOTS. Concluding remarks are made in Section 5.4.

5.1 Supporting Proportional QoS with Extra Offset Times on a Single Link

We now present an analysis on the use of an extra offset time to provide proportional QoS on a single link. Through the analysis, we try to address some fundamental yet important aspects of the extra offset time based proportional QoS scheme such as

1. What is the range of the achievable proportional ratio by using an extra offset time?

2. What is the achievable proportional ratio for a given extra offset time?

The analysis, which is valid on the first outgoing link, is not applicable in the edge-to-edge context. However, it provides insights to facilitate the understanding of our proposed extra offset time based edge-to-edge proportional QoS model described in Section 5.2. Specifically, the potential of using an extra offset time to provide proportional QoS can be examined in terms of whether a wide range of achievable proportional ratios can be delivered. Further, the study on what could affect the offset time needed to achieve a predefined proportional ratio can be carried out. We discuss these via numerical results of our analysis here.

Our analysis is based on the results in [28], [68], and [69]. In the single link model, we consider a node with no FDL buffer and k wavelengths on the link. The service differentiation is achieved via assigning different initial offset times, $T_{off}\{c_m\}$ to different classes of traffic for all N classes, where c_m ($m = 1, 2, \dots, N$) is the class type m . The lowest priority class is assigned a basic offset time denoted by T_{basic} , and a higher priority class is assigned an extra offset time denoted by $Tx\{c_m\}$ on top of T_{basic} . Bursts of a traffic class arrive according to a Poisson distribution with rate λ_m , and bursts receive service at an exponentially distributed service rate with mean $\mu = 1/L$, where L is the mean burst length. We also assume that the conservation law holds as in [28].

We consider a special case of the proportional model where the proportional loss ratio is defined between two neighbouring classes of traffic. That is, for any class m with dropping probability P_m , the proportional ratio for class m over class $m + 1$, denoted by $R_{m,m+1}$ is given by

$$R_{m,m+1} = \frac{P_m}{P_{m+1}} \quad m = 1, \dots, N - 1 \quad (5.2)$$

5.1.1 Achievable Proportional Ratio Range - Two Classes

In this section, we derive the achievable proportional loss ratio (or proportional ratio) range by using an extra offset time for service differentiation on a single link. Consider two classes where class 2 traffic has higher priority over class 1 traffic. The initial offset times used by class 1 and class 2 traffic are $T_{off}\{c_1\} = T_{basic}$, $T_{off}\{c_2\} = T_{basic} + Tx\{c_2\}$, respectively. From Eq. (5.2), we need to find out their respective dropping probabilities P_1 and P_2 in order to determine the proportional ratio of class 1 over class 2 traffic,

$$R_{1,2} = \frac{P_1}{P_2} \quad (5.3)$$

It has been shown in [28] that the dropping probability of a classless OBS system is given by the Erlang loss formula ($M/M/k/k$)

$$B(\rho, k) = \frac{\rho^k/k!}{\sum_{j=0}^k \rho^j/j!} \quad (5.4)$$

where ρ is the offered load given by λ/μ and k is the number of wavelengths.

The average overall dropping probability of an OBS system with N classes, according to the conservation law [28], is given by

$$\rho_{all}P_{all} = \sum_{j=1}^N \rho_j P_j \quad (5.5)$$

where P_{all} is the overall dropping probability, $\rho_j = \lambda_j/\mu$ and $\rho_{all} = \sum_{j=1}^N \rho_j$. In this case $N = 2$.

To find out the dropping probability of class 1 and class 2 traffic, intuitively there are two different scenarios to be considered. One is when complete isolation is achieved, when class 2 is not blocked by class 1 traffic. The other is when there is no complete isolation between the

two classes of traffic, i.e., a high priority burst can still be blocked by a lower priority burst because the extra offset time is insufficient to totally separate the two classes of traffic. We define the range of the achievable proportional ratio of a two-class case by $R_{1,2}^L \leq R_{1,2} \leq R_{1,2}^U$, where $R_{1,2}^L$ is the lower bound and $R_{1,2}^U$ is the upper bound. We consider the case of complete isolation in deriving $R_{1,2}^U$ due to the following Claim 1, while $R_{1,2}^L$ is derived using Claim 2.

Claim 1: The upper bound of the proportional ratio, $R_{1,2}^U$ (the maximum achievable proportional loss ratio), is determined by considering the case of complete isolation.

Proof: When complete isolation is achieved, class 2 bursts always “wins” when it comes to wavelength reservation. Therefore, the class 1 traffic encounters the highest dropping probability. Any further increase in the offset time after complete isolation is achieved will not affect the dropping probabilities of both the classes. ■

Since class 2 traffic is not blocked by any lower priority traffic, its blocking probability with complete isolation, denoted by P_2^{iso} , can be found by using the Erlang loss formula given by Eq. (5.4) and by considering the traffic from class 2 only, i.e.,

$$P_2^{iso} = B(\rho_2, k) = \frac{\rho_2^k/k!}{\sum_{j=0}^k \rho_2^j/j!} \quad (5.6)$$

The overall dropping probability is obtained by using the Erlang loss formula

$$P_{all} = B(\rho_{all}, k) = \frac{\rho_{all}^k/k!}{\sum_{j=0}^k \rho_{all}^j/j!} \quad (5.7)$$

Using Eq. (5.5), (5.6), and (5.7) and letting $P_2 = P_2^{iso}$, the dropping probability of class 1 traffic for the case of complete isolation, denoted by P_1^{iso} , is given by

$$P_1^{iso} = (\rho_{all}P_{all} - \rho_2P_2^{iso})/\rho_1 \quad (5.8)$$

Considering Claim 1, let $P_1=P_1^{iso}$, $P_2=P_2^{iso}$ for the case of complete isolation, $R_{1,2}^U$, is given by

$$R_{1,2}^U = \frac{P_1^{iso}}{P_2^{iso}} = (\rho_{all}P_{all}/P_2^{iso} - \rho_2)/\rho_1 \quad (5.9)$$

Claim 2: The lower bound of the proportional ratio, $R_{1,2}^L$, is one.

Proof: The lower bound of the proportional ratio is found by considering the case when no extra offset time is used, i.e. both classes of traffic use the same offset time as in the case of classless OBS. Therefore, both classes are treated equally during wavelength reservation and they have the same dropping performance. ■

5.1.2 Achievable Proportional Ratio Range - Arbitrary Number of Classes

Consider the general case of N traffic classes where apart from the lowest priority traffic, every higher priority traffic receives a larger extra offset time in increasing priority order, i.e. $Tx\{c_2\} < Tx\{c_3\} < \dots < Tx\{c_N\}$.

As the lower bound on the achievable proportional ratio is still one, we are interested in finding the upper bound of the achievable proportional ratio $R_{m,m+1}^U$. When there are more than two classes of traffic, say 3 classes, where class 1 traffic has the lowest priority and class 3 has the highest priority. In this case, the dropping probability of class 2 traffic P_2 will affect the proportional loss ratio between class 2 and class 3, $R_{2,3}$, as well as between class 1 and class 2, $R_{1,2}$. From the offset time perspective, if class 3 traffic is completely isolated from class 2 traffic, P_3 is at its lowest value. However, $R_{2,3}$ might not be the highest as it depends on P_2 as well and this is related to the extra offset time used to differentiate class 2 traffic from class 1 traffic. There are two different cases to be considered:

1. Class 2 traffic is completely isolated from class 1 traffic, therefore, class 2 traffic is not blocked by any class 1 traffic.
2. Class 2 traffic is not completely isolated from class 1 traffic, therefore, class 2 traffic might be blocked by class 1 traffic.

Intuitively, the dropping probability of class 2 traffic in case 2 is greater than that in case 1. Therefore $R_{2,3}$ is greater in case 2 than in case 1 (dropping probability of class 3 traffic is independent of these since it is completely isolated from class 2 traffic and therefore class 1 traffic as well). According to the conservation law for the OBS system, when P_2 increases (P_3 remains the same), P_1 decreases. This implies that $R_{1,2}$ decreases as $R_{2,3}$ increases. This means that the range of the achievable proportional ratio for two particular classes can be enlarged at the expense of a smaller range on the achievable proportional ratio for the other classes. A more detailed study is needed in order to determine the range of the achievable proportional loss ratio $R_{m,m+1}^U$ in a multi class environment.

5.1.3 Achievable Proportional Ratio for a Given Offset Time

Here, we would like to determine the achievable proportional loss ratio for a given offset time with two classes of traffic. Under the extra offset time based service differentiation method, when the extra offset time assigned to the higher priority traffic is not sufficiently large, complete isolation of traffic cannot be achieved. Therefore, to find the dropping probability of class 2 traffic, not only the offered load ρ_2 but also a fraction of the carried traffic of the low priority class 1 traffic that interferes with class 2 traffic needs to be considered. This fraction of class 1 carried traffic is, according to [68], given by

$$y_1(\tau) = \rho_1(1 - P_1)(1 - F_1(\tau)) \quad (5.10)$$

where τ is the class 2 traffic offset time, and $F_1(\tau)$ is the distribution function of class 1 bursts whose length L is not longer than τ , given by

$$F(\tau) = P[L \leq \tau] = 1 - P[L > \tau] = 1 - \exp^{(-1/L)\tau} \quad (5.11)$$

Note that if τ is big enough, the following steps would lead to the calculation of the achievable proportional ratio for the case with complete isolation.

Knowing the carried traffic from class 1, the dropping probability of high priority class 2 traffic P'_2 is approximated by

$$P'_2 = B(\rho_2 + y_1(\tau), k) \quad (5.12)$$

The dropping probability of low priority class 1 traffic, P'_1 can be obtained using the conservation law, i.e.,

$$P'_1 = (\rho_{all}P_{all} - \rho_2P'_2)/\rho_1 \quad (5.13)$$

As there is a mutual dependency between P'_1 and P'_2 according to Eq. (5.10), (5.12), and (5.13), an iterative solution is needed as suggested in [68, 69]. Let P''_1 and P''_2 be the final results after some iterations where a given precision criterion is satisfied. The achievable proportional loss ratio for the general case (either with or without complete isolation) denoted by $R''_{1,2}$, is given by

$$R''_{1,2} = \frac{P''_1}{P''_2} = (\rho_{all}P_{all}/P''_2 - \rho_2)/\rho_1 \quad (5.14)$$

It has been pointed out in [69] that it is difficult and computationally complex to find the reasonable offset time for a required burst dropping probability as an iterative method is

used to determine the burst dropping probability for a given offset time. Therefore, it is also not trivial to find the offset times needed for different classes of traffic in order to achieve a predefined proportional ratio in dropping probability.

For an environment with multiple classes of traffic, all interferences from a class w lower priority traffic on a class r higher priority traffic ($1 \leq w < r, 1 < r \leq N$) need to be considered. The above derivation of the achievable proportional ratio can be extended for any two adjacent classes of traffic where interference from more than one class of traffic is considered. The expression for quantifying the carried traffic from all the lower priority traffic which affect the higher priority traffic can be found in [68].

5.1.4 Numerical Results

We now present some numerical results of our analysis. It is used to provide some insights into using an extra offset time to support proportional QoS. We examine the achievable range of proportional loss ratio as well as outlining different scenarios where the offset times needed to achieve a predefined proportional loss ratio are different. Here, two classes of traffic are considered where class 1 is the low priority traffic and class 2 is the high priority traffic. The mean burst length L is assumed to be 1 unit and for the case without complete isolation (i.e., for the calculation of $R''_{1,2}$), an extra offset time of 1 unit is used for the class 2 traffic. The total traffic intensity in the range of 0.1 to 1 is considered.

Impact of the Number of Wavelengths on Maximum Achievable Proportional Ratio

We first show how the upper bound of the proportional ratio range is affected by the number of wavelengths. Recall that the maximum achievable proportional ratio for the case with

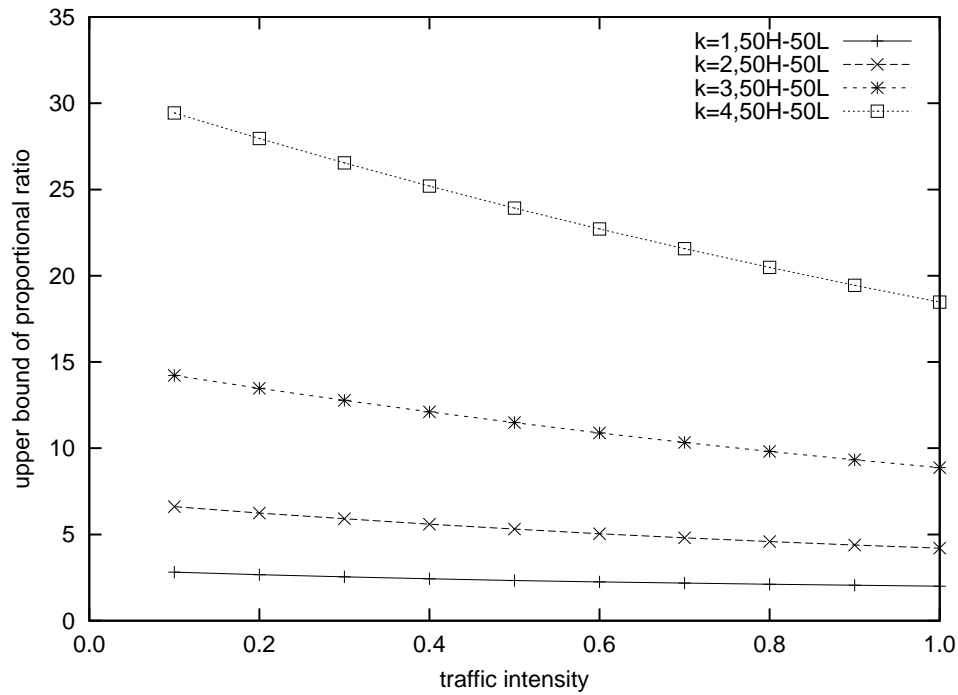


Figure 5.1: Upper bound of the achievable proportional ratio, $R_{1,2}^U$ (with complete isolation) for traffic composition 50H-50L.

complete isolation is given by Eq. (5.9). Figure 5.1 shows that $R_{1,2}^U$ increases with the number of wavelengths¹. Thus, as the number of wavelengths in the WDM network is large, the offset time scheme is able to provide a large range of proportional ratios.

Impact of the Traffic Composition on the Maximum Achievable Proportional Ratio

We now examine how traffic composition affects the upper bound of proportional ratio range. With higher proportion of low priority traffic to high priority traffic as shown in Figure 5.2 compared to Figure 5.1, $R_{1,2}^U$ in Figure 5.2 is much higher than that in Figure 5.1. Therefore, as ρ_1/ρ_2 increases, the range of the proportional ratio expands as well. This is beneficial to

¹50H-50L refers to the case where 50% of the traffic are high priority traffic while the other 50% of the traffic are low priority traffic.

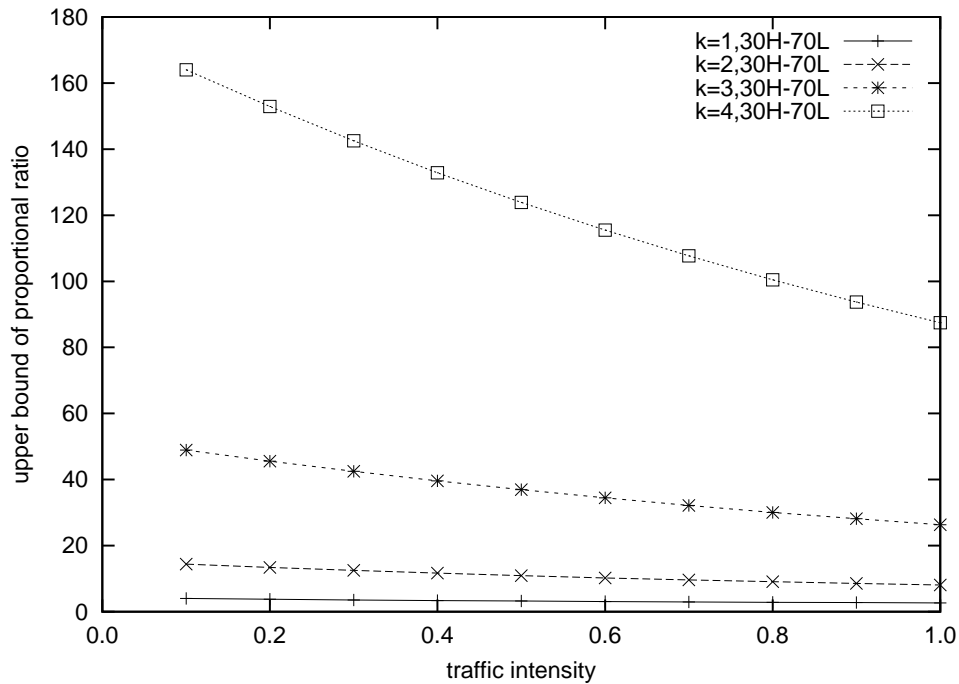


Figure 5.2: Upper bound of the achievable proportional ratio, $R_{1,2}^U$ (with complete isolation) for traffic composition 30H-70L.

the real network case where lower priority traffic (e.g. best effort traffic) is dominant as a large range of proportional ratio for service differentiation can be provided.

Impact of the Traffic Composition on Offset Time

The achievable proportional ratio for the case without complete isolation is given by Eq. (5.14). Figure 5.3 shows that with increasing traffic intensity, the achievable proportional ratio for the case of 50H-50L decreases. Figure 5.4 shows a similar trend for the case of 30H-70L. Therefore, in order to maintain the same proportional ratio (as with lower traffic intensity), the extra offset time needed by the high priority burst increases with increasing traffic intensity. This is more pronounced when the number of wavelengths increases as Figures 5.3 and 5.4 show. Further, it is observed that the achievable proportional ratio (with the same amount of extra offset time, i.e. 1 unit) is higher with a traffic composition of

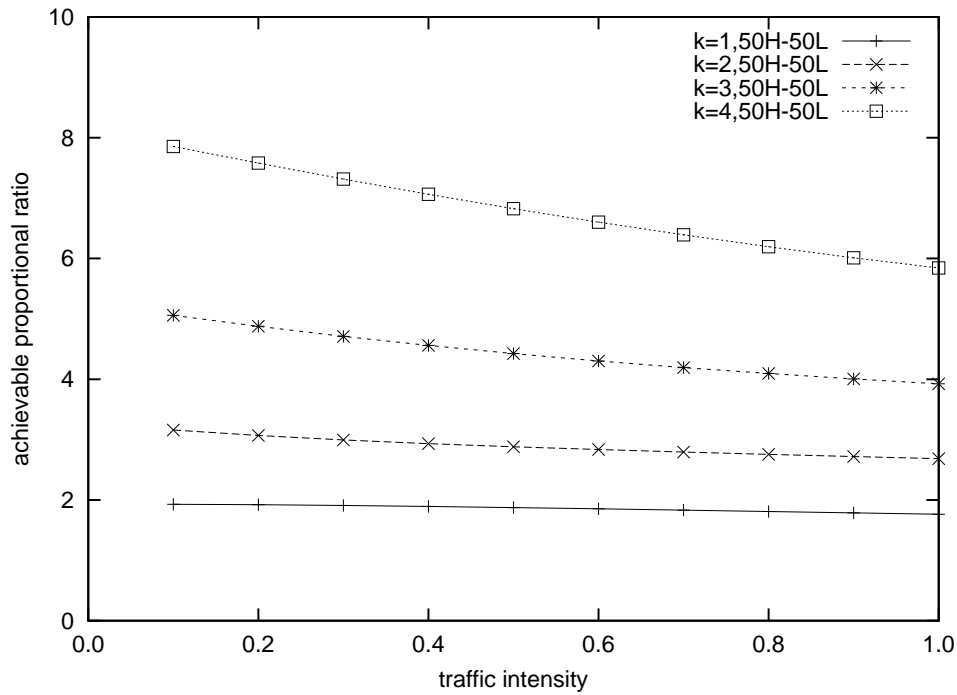


Figure 5.3: Achievable proportional ratio, $R''_{1,2}$ (without complete isolation) for traffic composition 50H-50L.

30H-70L than with a 50H-50L traffic composition. This means that the offset time needed in order to achieve a certain proportional ratio is shorter with the former traffic composition. Indeed, for $k = 4$ $\rho = 0.2$, and the same offset time used by class 2 traffic, the achievable proportional ratio, $R''_{1,2}$ in Figure 5.4 is about 13 while it is about 7.6 in Figure 5.3. Thus, the offset time needed to achieve a predefined proportional ratio decreases with increasing proportion of low-priority traffic.

5.2 Proposed FOTS Method

As discussed in Section 5.1, it has been pointed out in [69] that it is difficult to determine the offset time needed to achieve a certain burst dropping probability. It is therefore non trivial to find the offset times needed for different traffic classes to achieve a predefined proportional

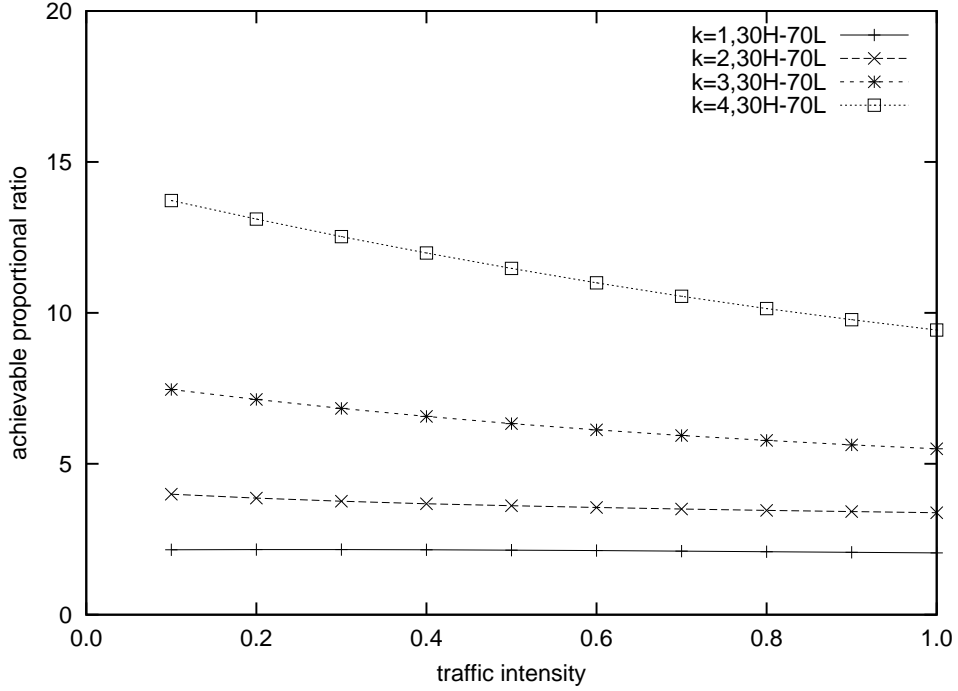


Figure 5.4: Achievable proportional ratio, $R''_{1,2}$ (without complete isolation) for traffic composition 30H-70L.

ratio on a single link, and tougher still to achieve an edge-to-edge proportional ratio for different node pairs with different hop counts. We are thus motivated to find a more practical method to realize edge-to-edge proportional QoS. We present the proposed FOTS method in this section as one such practical method.

5.2.1 Overview of FOTS

The proposed FOTS method exploits the adjustable offset time feature at the OBS edge nodes to realize edge-to-edge proportional QoS provisioning. This ensures that no complex processing is incurred at the core nodes. What FOTS does is that it simply selects the offset time dynamically at ingress nodes based on feedback information collected from egress nodes and from the probe packets sent to the core nodes. Label switched paths (LSPs) are assumed to be set up through extension of *multi-protocol label switching* (MPLS) capabilities. Initially,

different classes of traffic use a set of initial offset times. A probe packet is sent periodically along each LSP to collect the link scheduling states (explained later) with respect to various offset times, both larger and smaller than the current offset time, called candidate offset times. Every egress node keeps track (in a burst arrival counter) of the number of bursts of different classes that arrive on the LSPs from different ingress nodes and feeds this information back to the ingress node via the probe packet at the end of the traffic measurement period. The burst dropping probability with respect to an ingress-egress node pair and to a traffic class can then be computed with this information, from which FOTS decides on the new set of initial offset times for the different traffic classes chosen from their respective candidate offset times to meet the predefined proportional ratios.

As shown in Figure 5.5, FOTS operates in cycles, each cycle being defined as one traffic measurement period. A probe packet is sent on each LSP at the end of a traffic measurement period. The probe packet collects the link state of every node along the LSP. Upon reaching the egress node, it records the total number of burst arrivals counted by the burst arrival counter at the egress node and is returned to the ingress node through the LSP in the reverse order. Figure 5.6 illustrates an example of the probe packet format with important fields catered for link state collection that include LSP label, class, burst size, initial offset time, hop number, table size, and its associated flag bits which gives an indication of the link availability if the burst assigned with it different offset times. We assume that the network traffic is relatively stable and does not change abruptly during a traffic measurement period [76] and the aggregate traffic on links changes very slowly as observed in [79] and [80]. Further, we also assume that the round-trip propagation time of a probe packet is short relative to the traffic measurement period.

The proposed FOTS has the following attractive features:

- The traffic loading patterns are captured using online link states.

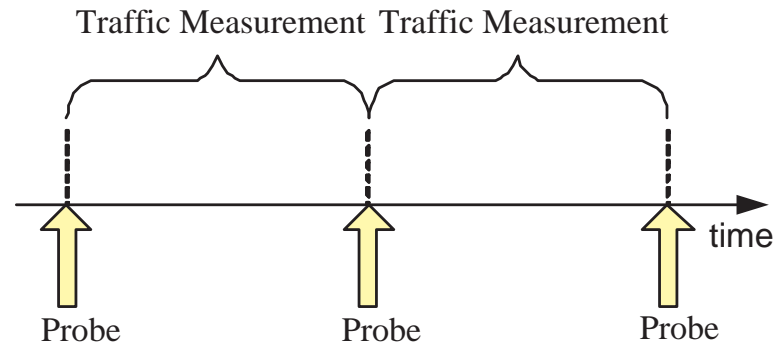


Figure 5.5: Illustration of FOTS with probing for traffic measurement collection and traffic measurement period on time axis.

LSP Label	Class	Burst Size	Initial offset time	Hop number	Table size	Flag bits
-----------	-------	------------	---------------------	------------	------------	-----------

Figure 5.6: Probe packet format for link state collection

- No intentional dropping of bursts is needed, thus, resource utilization is improved.
- Signalling overhead is small as the new set of offset times selected is used for a sufficiently longer time before the next probe packet is sent.
- Intelligent decisions are taken at the edge nodes rather than at the core nodes, relieving the core nodes of the processing and algorithmic burden.
- Fairness in terms of the proportional ratio achieved for node pairs with various hop lengths is maintained.

In the following, we illustrate the operation of FOTS with respect to one traffic measurement period and for a pair of nodes.

5.2.2 Link State Collection

Under FOTS, link state collection is done by a probe packet on every LSP. The link state at time t and offset time τ indicates the availability of a free wavelength on the link at time $t+\tau$. At the end of a traffic measurement period, a probe packet carrying an offset time table (Table 5.1) probes for the link availability with various initial offset times. The entries in Table 5.1 are offset times for the node pair v for class 2 traffic and the corresponding link availability status within a traffic measurement period.

The probe packet searches for the availability of a free wavelength at the various candidate offset times. These are set at larger and smaller values than $T_{off}\{v, c_2\}$ given by $T_{off}\{v, c_2\} + sz$ where $0 \leq |s| \leq M$, i.e. $T_{off}\{v, c_2\} + z, T_{off}\{v, c_2\} + 2z, \dots, T_{off}\{v, c_2\} + Mz$ and $T_{off}\{v, c_2\} - z, T_{off}\{v, c_2\} - 2z, \dots, T_{off}\{v, c_2\} - Mz$ respectively where M and z are predefined. M and z have to be chosen carefully so that they provide a suitable range of offset times for FOTS to adapt to the changing traffic load pattern quickly enough while not imposing too much overheads. The availability column of the probe table is updated (indicated as 1 if the link is available and zero otherwise) as the probe packets traverse the LSP. Obviously, entries in the offset time table with no free wavelength available need not be checked on the subsequent links in an LSP. The probing action continues until the probe packet reaches the egress node where it is returned on the reverse path to the ingress node. From the offset time table, the ingress node is able to find out among those candidate offset times, which ones allow the burst to successfully traverse along the LSP (the offset time entry with a link availability indicated by 1). We note that for increased accuracy of the link state information collected, several probe packets can be sent continuously at the expense of higher signalling overhead.

Table 5.1: Offset time table carried by a probe packet.

Offset	Link Availability (1/0)
$T_{off}\{v, c_2\} - 2z$	0
$T_{off}\{v, c_2\} - z$	1
$T_{off}\{v, c_2\}$	1
$T_{off}\{v, c_2\} + z$	1
$T_{off}\{v, c_2\} + 2z$	1

5.2.3 Traffic Measurement

Traffic measurement in FOTS involves a simple process where a burst arrival counter at an egress node is incremented upon the arrival of a burst on an LSP for a given class of traffic and ingress node. This measurement (via updating the burst arrival counter) is done for the traffic measurement period, T_p , and is reset to zero when this information is sent back to the ingress node in a probe packet. Therefore, this traffic measurement reflects the dropping performance as a result of each offset time decision. Note that it can be the same probe packet which is used to collect the link states. An additional field has to be added to the probe packet shown in Figure 5.6 for this purpose. Alternatively, a different probe packet can be sent specifically for traffic measurement collection. As the total number of bursts sent is known to the ingress node, the ingress node is able to calculate the burst dropping performance of different ingress-egress node pairs belonging to different traffic classes with this information.

Notations:

$x(v, c_m)$: Total number of bursts of class c_m recorded at the ingress node for node pair v ;

$a_e(v, c_m)$: Total number of bursts that arrive at the egress node of class c_m for node pair v ;

The actions performed during a traffic measurement period are given below:

1. At the beginning of the measurement period:

- at the ingress node: Set $x(v, c_m)$ to zero.
- at the egress node: Set $a_e(v, c_m)$ to zero.

2. During the measurement period:

- When a burst of class c_m arrives at the ingress node:

$$x(v, c_m) = x(v, c_m) + 1,$$

- When a burst of class c_m arrives at the egress node:

$$a_e(v, c_m) = a_e(v, c_m) + 1,$$

3. At the end of the measurement period: Send the probe packet

- At the egress node: Upon receiving the probe packet send it back to the ingress node with the value of $a_e(v, c_m)$.
- At the ingress node: Upon receiving back the probe packet, the mean burst dropping probability for the node pair v , denoted by $p(v, c_m)$ is calculated as follows:

$$p(v, c_m) = \frac{x(v, c_m) - a_e(v, c_m)}{x(v, c_m)} \quad (5.15)$$

5.2.4 Offset Time Selection

The offset time selection is done at the end of each traffic measurement period upon receiving back the probe packet sent from the egress node. Based on the link states (i.e., the offset time table) and the traffic measurements (i.e., the burst dropping probabilities) FOTS selects the offset time needed for all node pairs at the ingress node. We illustrate the procedure with respect to one pair of nodes v with two traffic classes. Let class m and class $m + 1$ traffic be the low priority and high priority traffic respectively. The initial offset time used is given by

$T_{off}\{v, c_m\}$. The predefined proportional ratio between class m and class $m + 1$ is given by $R_{m,m+1}^d$, where $R_{m,m+1}^d > 1$. FOTS first computes the achieved proportional ratio between the two classes of traffic within a traffic measurement period, denoted by $R'_{m,m+1}$ and compare $R'_{m,m+1}$ with $R_{m,m+1}^d$. If the predefined proportional ratio has not been achieved, a new offset time is selected. A higher offset time ($> T_{off}\{v, c_{m+1}\}$) is considered if $R'_{m,m+1} < R_{m,m+1}^d$, and a lower offset time ($< T_{off}\{v, c_{m+1}\}$) is considered if $R'_{m,m+1} > R_{m,m+1}^d$. The available offset time with the smallest $|s|$ will be selected accordingly as the new offset time. The details of the algorithm are given in Table 5.2.

5.2.5 Supporting More than Two Traffic Classes

FOTS can be extended to accommodate more than two classes of traffic. This is done by sending probe packets for each class of traffic where each probe packet carries an offset table inclusive of the current initial offset time used for the respective class. The probe packets belonging to the different classes of traffic take turns to transmit (the lowest priority burst uses the basic offset time). Upon receiving back the probe packet at the ingress node, the offset time selection process for the respective class is invoked while the offset times for the remaining traffic classes stay unchanged. For example, consider three classes of traffic denoted by class 1, class 2 and class 3. Class 1 is the lowest priority traffic and its offset time is not changed. At the end of the first traffic measurement period, class 2 transmits a probe packet on every LSP. The probe packet carries out the link state collection. The return probe packet carries with it the link states as well as the traffic measurements recorded at the egress node. Upon receiving the returned probe packet the offset time selection for class 2 traffic is invoked. At the second traffic measurement period, the new offset times are used for class 2 traffic while the offset time for all other classes of traffic remain unchanged. At the end of the second traffic measurement period, class 3 sends a probe packet on every LSP carrying

Table 5.2: Offset Time Selection in LSOS

Input: $R_{m,m+1}^d$, $p(v, c_m)$, $p(v, c_{m+1})$, offset time table

Step 1: $T_{offset} \leftarrow T_{off}\{v, c_{m+1}\};$

Step 2: Compute proportional ratio achieved in the current traffic measurement period:

$$R'_{m,m+1} = \frac{p(v, c_m)}{p(v, c_{m+1})} \quad (5.16)$$

Step 3: Select new available offset time:

case 1:

if ($R'_{m,m+1} < R_{m,m+1}^d$)

for $s = 1, 2, \dots, M$ **do**

 if $T_{off}\{v, c_{m+1}\} + sz$ is available (i.e. link availability =1)

 then $T_{offset} \leftarrow T_{off}\{v, c_{m+1}\} + sz$; break;

endif

endif **endfor**

case 2:

if ($R'_{m,m+1} > R_{m,m+1}^d$)

for $s = -1, -2, \dots, -M$ **do**

 if $T_{off}\{v, c_{m+1}\} + sz$ is available (i.e. link availability =1)

 then $T_{offset} \leftarrow T_{off}\{v, c_{m+1}\} + sz$; break;

endif

endif **endfor**

Step 4: Update the initial offset time:

$T_{off}\{v, c_{m+1}\} \leftarrow T_{offset}$

out similar procedures and the offset time for class 3 traffic are adjusted at the beginning of the following traffic measurement period. A period of time is thus needed for the offset time selection process for the various traffic classes to settle down (i.e., proportional ratio is maintained with little changes in offset time needed). With increasing number of classes, the settling down time required will be longer.

5.2.6 Convergence and Stability Issues

The convergence and stability issues are some of the concerns in the FOTS method, just like in any other feedback based mechanism. In the context of achieving the predefined edge-to-edge proportional QoS, convergence issues are such as whether FOTS takes a long time to achieve the predefined proportional ratio, or it ever reaches the goal. Stability issues are such as whether the proportional ratio achieved is always observed to be the same as the predefined proportional ratio value over times. Although our assumption of a relatively stable network traffic generally holds, however, due to the dynamic behaviour of the arrival traffic, the predefined proportional ratio will not be achieved all the time. Despite this generally observed deviation from the desired goal for most feedback based mechanism, the network usually reaches a “dynamic” equilibrium state with oscillation around the equilibrium state. Here, the time needed to reach this state and the magnitude of the oscillation gives a good measure of the convergence [87]. In the context of FOTS, the simulation results in Section 5.3 show that after sometimes (where different classes of traffic take times to adjust their respective offset times), the average proportional ratio achieved will converge and oscillate (in small magnitude) around the predefined proportional ratio, and similar performance is observed over time.

5.3 Performance Study

The performance of FOTS has been studied through simulation. A 10-node randomly generated topology with edge-to-edge hop count ranging from 1 to 5 is used. Explicit routing with fixed shortest path is used. Each link on the network has 8 channels for data traffic and 1 channel for control traffic. The transmission capacity of each channel is approximately 10 Gpbs. Burst arrival follows a Poisson process and burst lengths are exponential distributed. The mean burst length used is 16 μsec and the control packet processing time is 2 μsec . Three classes of traffic are used. Class 1 traffic has the lowest priority and class 3 traffic has the highest priority. In the simulations, we set $M = 5$ and $z = 1 \mu sec$. We note that a large value of M incurs high signalling and computational overhead and a small value of M cannot predict the link state accurately. Therefore, $M = 5$ is a reasonable choice such that it is sufficient to support the 3 classes of traffic with the predefined proportional ratios.

The performance of FOTS has been studied with respect to the following:

- traffic measurement period T_p
- traffic proportion
- traffic load
- proportional ratio

For ease of exposition, we denote the traffic proportion as x-y-z corresponding to the percentage of traffic classes 3-2-1. We consider two different traffic proportions given by 30-30-40 and 20-20-60. Two burst arrival rates of 0.1 and 0.2 (bursts/ μsec) are used that correspond to approximately 1% and 10% overall dropping probability in the simulation. The predefined proportional ratios for class 1 over class 2 traffic and class 2 over class 3 traffic are denoted by $R_{1,2}^d$ and $R_{2,3}^d$, respectively. We plot the proportional ratio achieved over successive probe

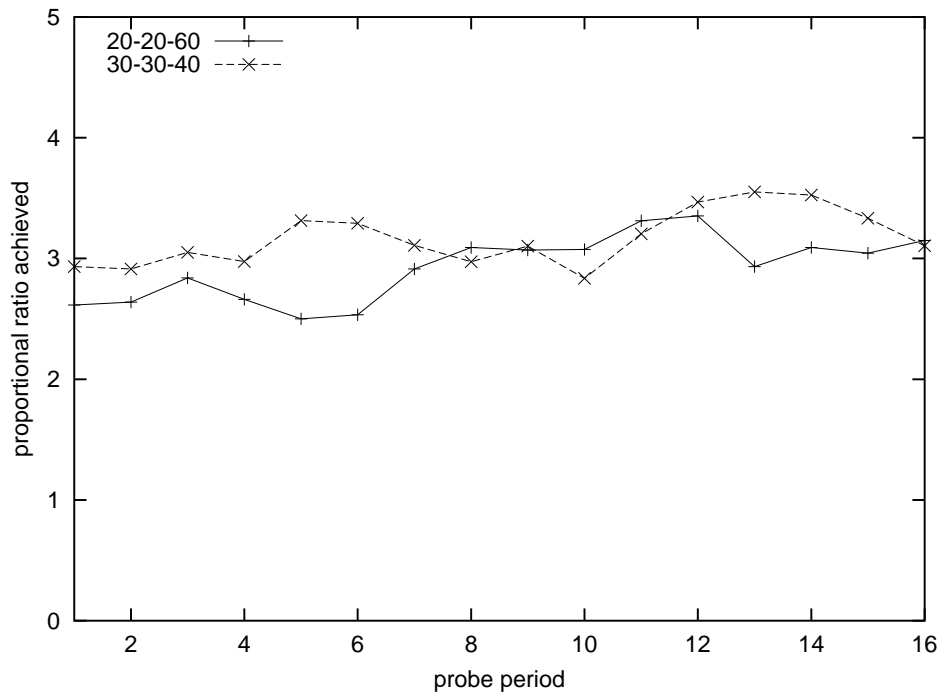


Figure 5.7: Proportional ratio achieved between class 1 and class 2, with $R_{1,2}^d = 3$, $T_p = 50$ msec and arrival rate of 0.1 bursts/ μ sec.

periods, each of length T_p . The proportional ratio achieved for various hop counts is also plotted to show the effectiveness of FOTS in the ensuring fairness in the proportional ratio achieved.

Figure 5.7 shows that with a burst arrival rate of 0.1, $R_{1,2}^d = 3$ and $T_p = 50$ msec, both traffic proportions of 20-20-60 and 30-30-40 achieve the set proportional ratio between class 1 and class 2 traffic. More stable results are observed with $T_p = 100$ msec as shown in Figure 5.8. This is because with a longer traffic measurement period, information of better representation of the network state is obtained resulting in a better offset time decision. For class 2 traffic over class 3 traffic with $R_{2,3}^d = 5$, the proportional ratio achieved for $T_p = 50$ msec as shown in Figure 5.9 and for $T_p = 100$ msec as shown in Figure 5.10 are close to 5. Also the performance of FOTS with $T_p = 100$ msec is more stable.

The performance of FOTS for a higher burst arrival rate of 0.2 (resulting in 10% overall burst

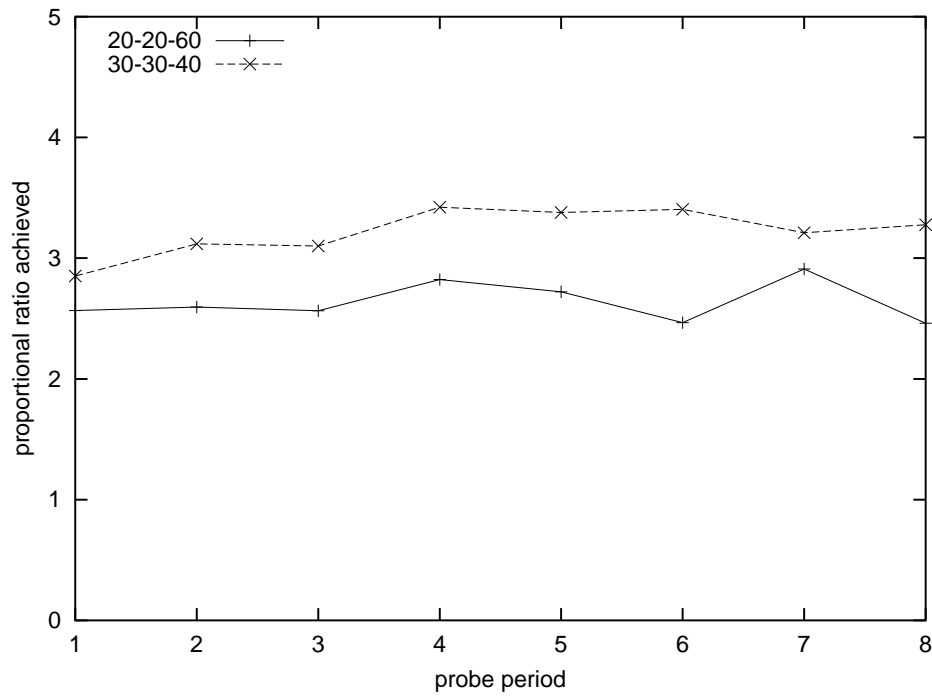


Figure 5.8: Proportional ratio achieved between class 1 and class 2, with $R_{1,2}^d = 3$, $T_p = 100$ msec and arrival rate of 0.1 bursts/ μ sec.

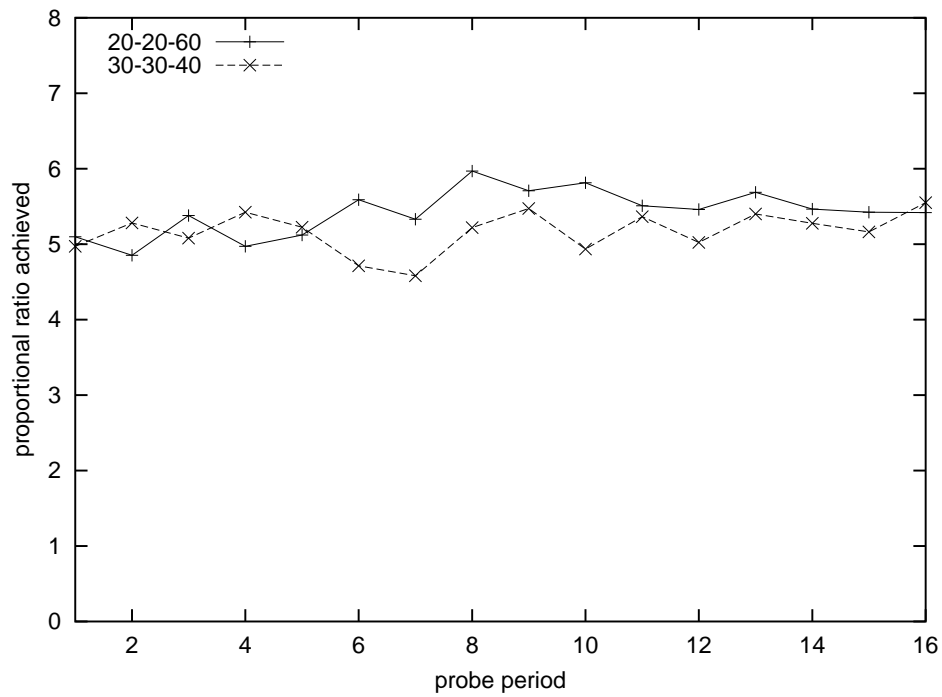


Figure 5.9: Proportional ratio achieved between class 2 and class 3, with $R_{1,2}^d = 3$, $T_p = 50$ msec and arrival rate of 0.1 bursts/ μ sec.

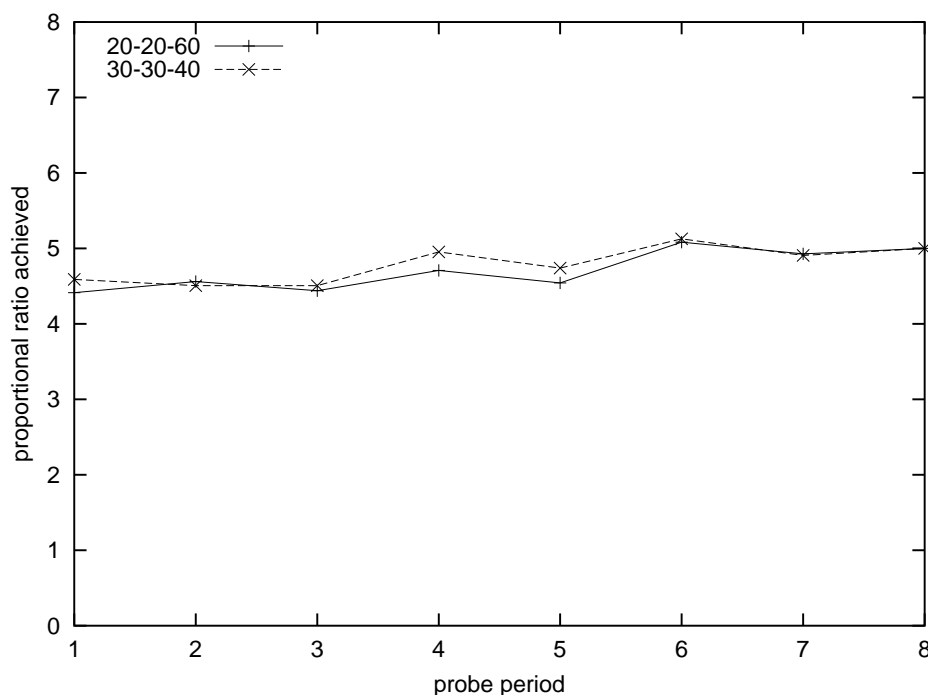


Figure 5.10: Proportional ratio achieved between class 2 and class 3, with $R_{1,2}^d = 3$, $T_p = 100$ msec and arrival rate of 0.1 bursts/ μ sec.

dropping probability) is also studied to show the effectiveness of FOTS under a busy network condition. For $R_{1,2}^d = 3$ and $T_p = 50$ msec (Figure 5.11) and for $T_p = 100$ msec (Figure 5.12), the proportional ratio achieved for class 1 over class 2 traffic is very close to 3. With $R_{2,3}^d = 5$ and $T_p = 50$ msec (Figure 5.13) and for $T_p = 100$ msec (Figure 5.14), the performance of FOTS is also good as the proportional ratio achieved for class 2 over class 3 traffic is close to 5. These results are also more stable than those for the lower arrival rate of 0.1 because the larger number of bursts lead to a more accurate offset time decision.

Next, the achieved proportional ratios for various hop counts are presented to show the effectiveness of FOTS in maintaining fairness among node pairs with different hop counts. For the lower arrival rate of 0.1, with $R_{1,2}^d = 3$, Figure 5.15 shows the proportional ratio achieved for various hop counts for different traffic proportions and different values of T_p . Node pairs of various hop counts are able to achieve the proportional ratio close to 3 for class

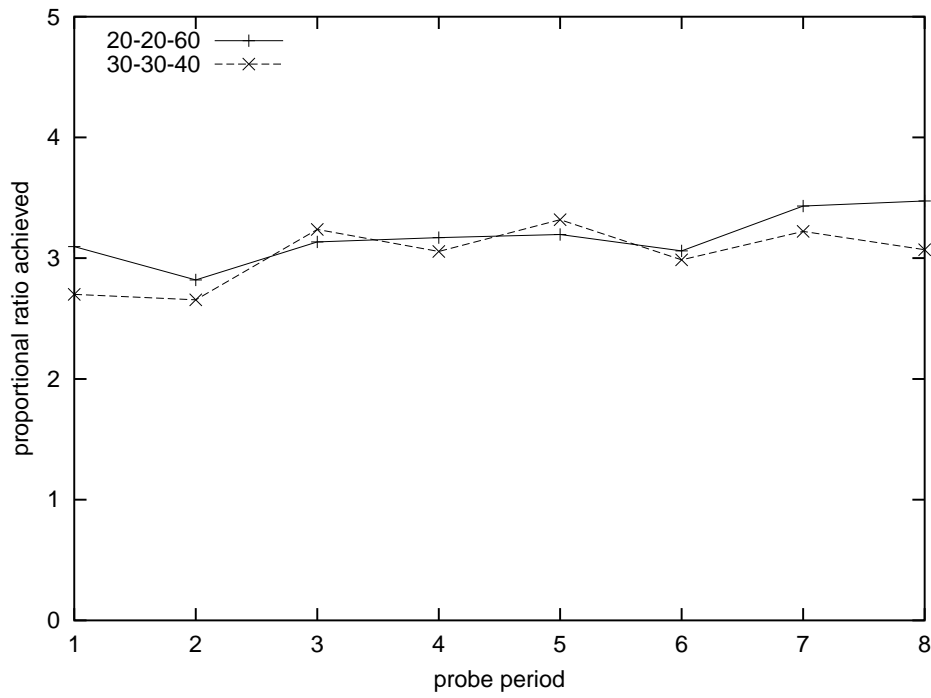


Figure 5.11: Proportional ratio achieved between class 1 and class 2, with $R_{1,2}^d = 3$, $T_p = 50$ msec and arrival rate of 0.2 bursts/ μ sec.

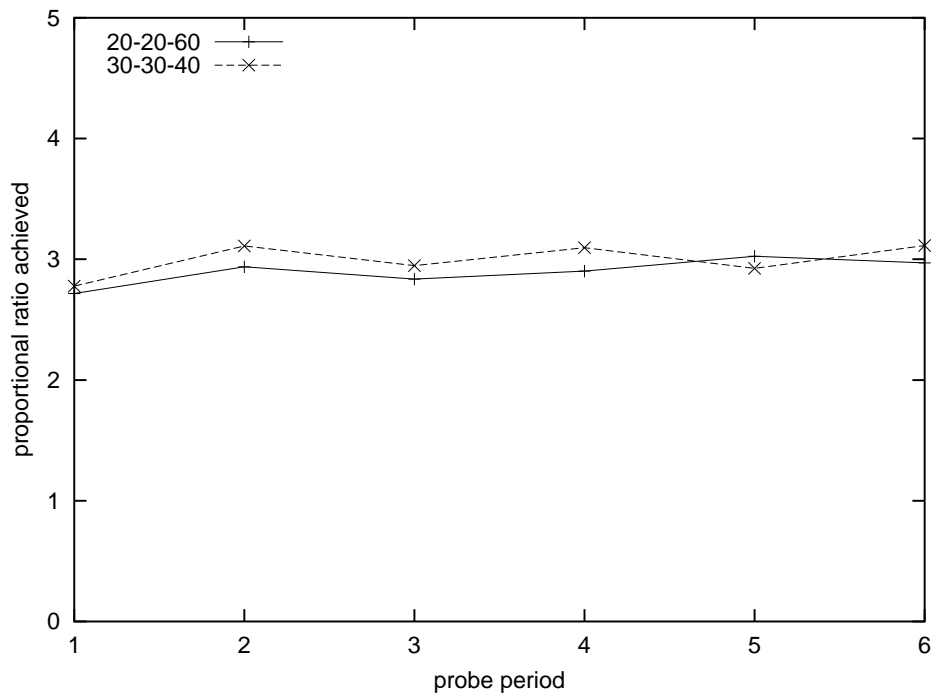


Figure 5.12: Proportional ratio achieved between class 1 and class 2, with $R_{1,2}^d = 3$, $T_p = 100$ msec and arrival rate of 0.2 bursts/ μ sec.

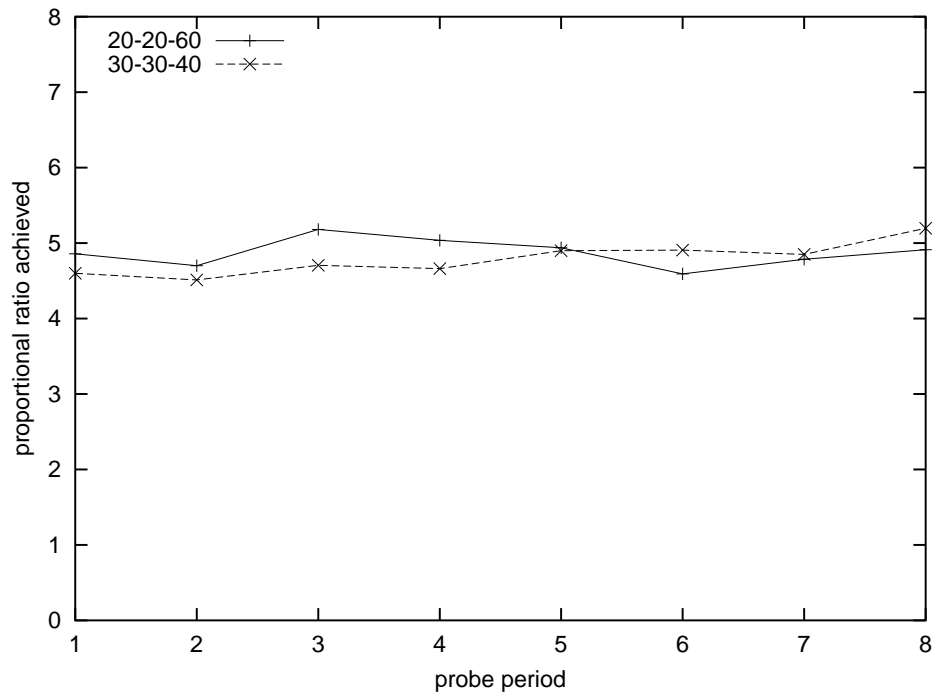


Figure 5.13: Proportional ratio achieved between class 2 and class 3, with $R_{2,3}^d = 5$, $T_p = 50$ msec and arrival rate of 0.2 bursts/ μ sec.

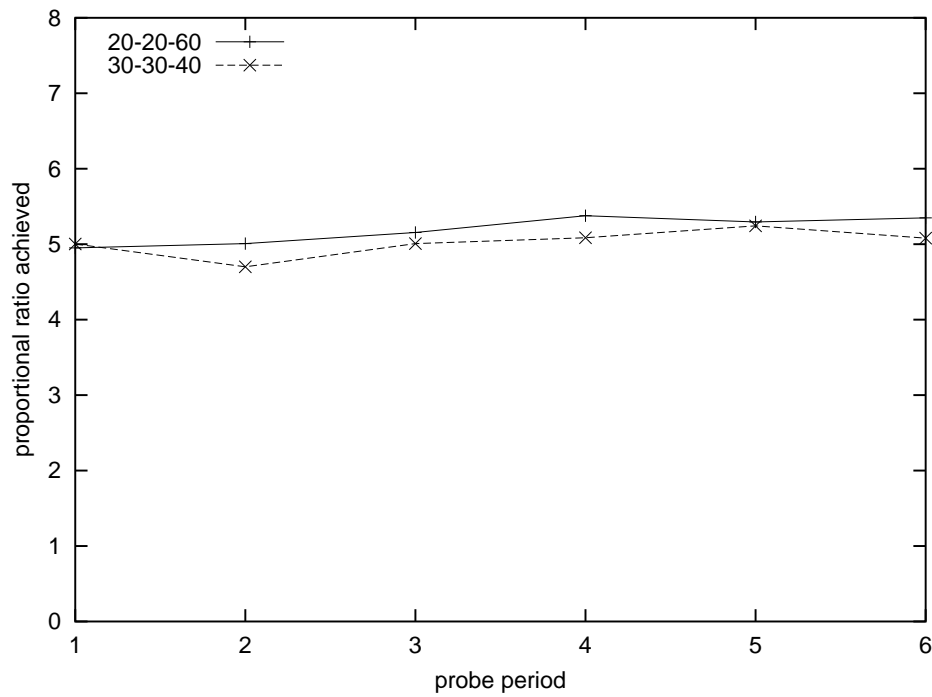


Figure 5.14: Proportional ratio achieved between class 2 and class 3, with $R_{2,3}^d = 5$, $T_p = 100$ msec and arrival rate of 0.2 bursts/ μ sec.

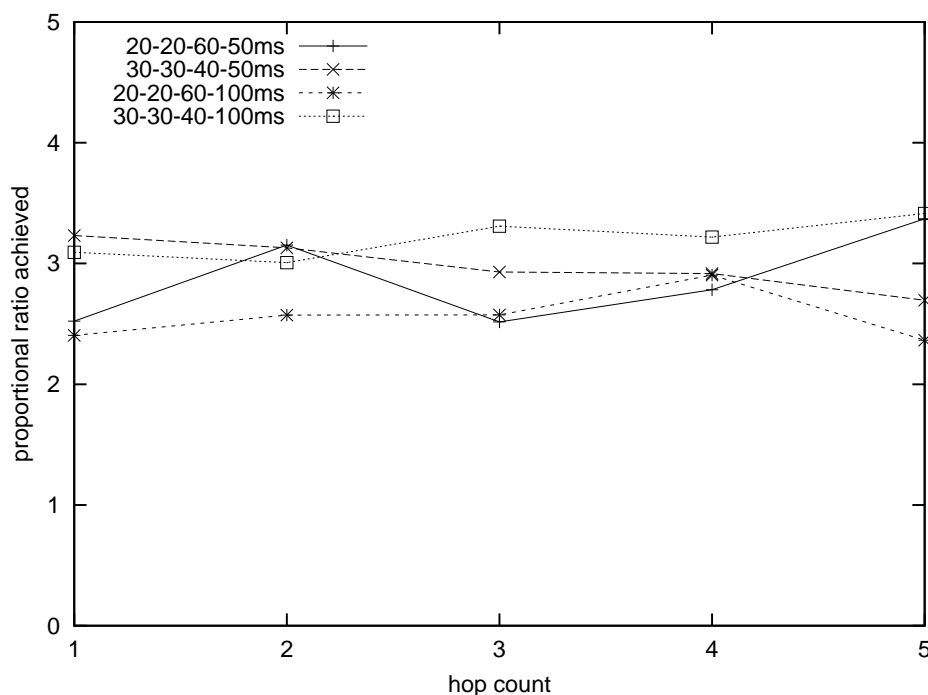


Figure 5.15: Proportional ratio achieved between class 1 and class 2, with $R_{1,2}^d = 3$, and arrival rate of 0.1 *bursts*/ μ sec.

1 over class 2 traffic. Similar results are achieved for a higher arrival rate of 0.2 as shown in Figure 5.16. Similarly, Figures 5.17 and 5.18 show that the proportional ratio achieved for class 2 over class 3 traffic for node pairs with different hop counts is close to 5 for burst arrival rates of 0.1 and 0.2, respectively.

With the arrival rate of 0.1, $R_{1,2}^d = 3$, the offset times needed for different traffic proportions and different traffic measurement periods are shown in Figure 5.19. The average offset time needed in the case of 20-20-60 traffic proportion is lower than that needed in the case of 30-30-40 traffic proportion. This is because the former case has a higher traffic proportion for the lowest priority traffic compared to the latter case. This is in line with the discussion in Section 5.1.4. Figure 5.20 shows that with an arrival rate of 0.2, the difference in terms of offset times for the two traffic proportions has diminished. For class 2 over class 3 traffic, with $R_{2,3}^d = 5$, the average offset time for node pairs with various hop counts are presented in

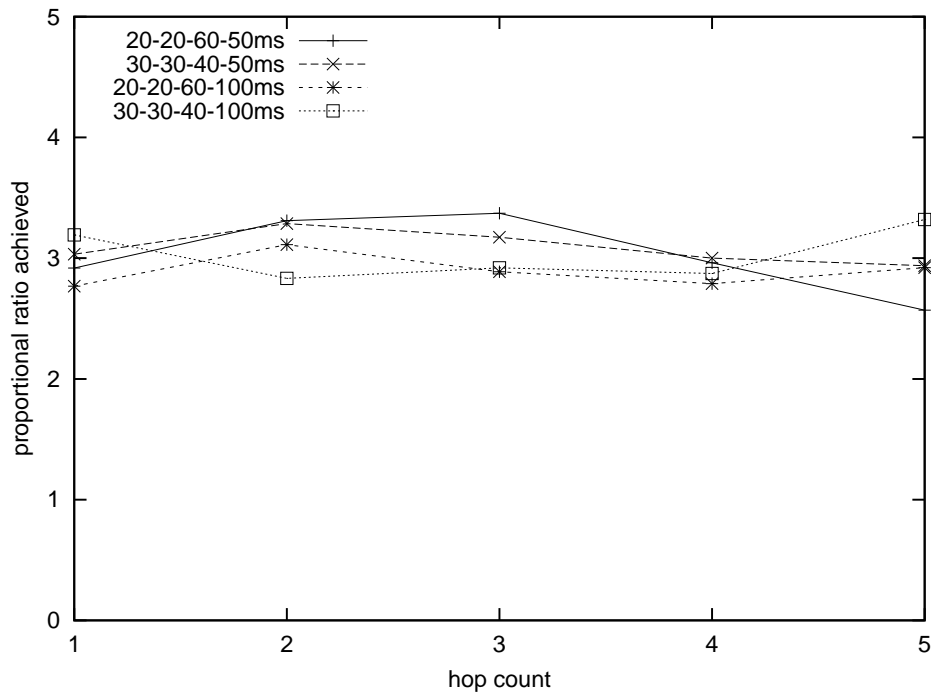


Figure 5.16: Proportional ratio achieved between class 1 and class 2, with $R_{1,2}^d = 3$, and arrival rate of $0.2 \text{ bursts}/\mu\text{sec}$.

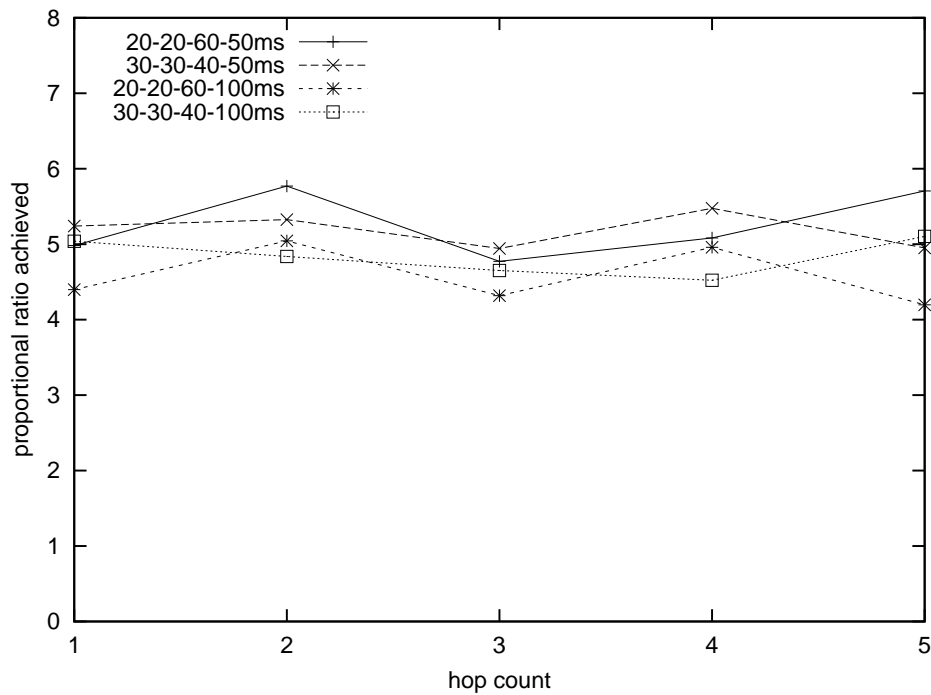


Figure 5.17: Proportional ratio achieved between class 2 and class 3, with $R_{2,3}^d = 5$, and arrival rate of $0.1 \text{ bursts}/\mu\text{sec}$.

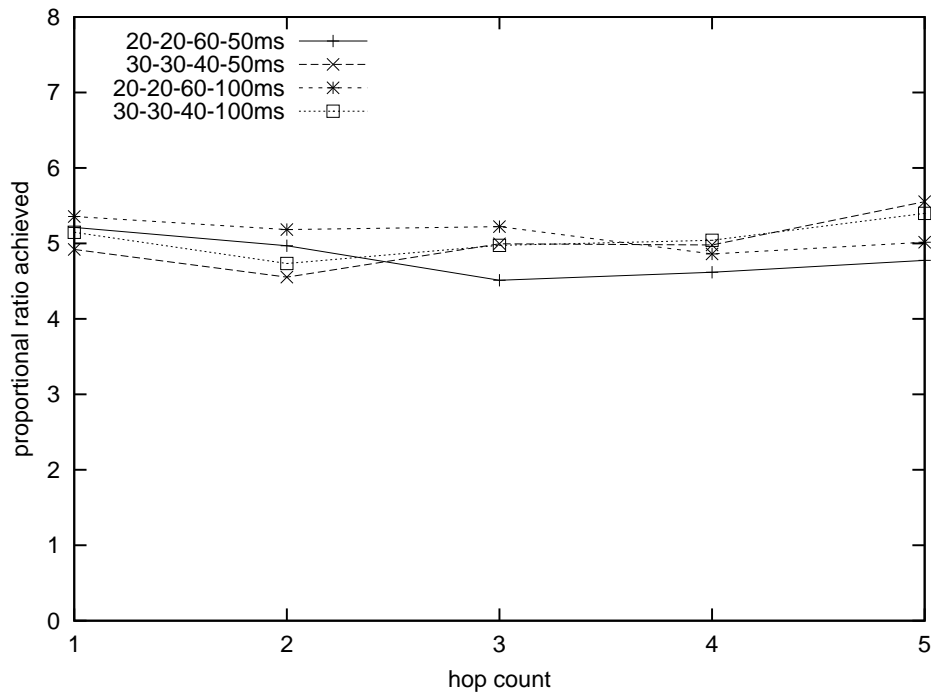


Figure 5.18: Proportional ratio achieved between class 2 and class 3, with $R_{2,3}^d = 5$, and arrival rate of $0.2 \text{ bursts}/\mu\text{sec}$.

Figure 5.21 for the arrival rate of 0.1, and in Figure 5.22 for the arrival rate of 0.2. Similar trends are observed as in the case of class 1 over class 2 traffic.

For larger network topology, FOTS is likely to take a longer time in converging to a stable value with respect to the predefined proportional QoS. This is due to the fact that with more node pairs involved in a larger topology, the case where adjustment of one pair affecting the performance (hence adjustment of offset time) of the other node pairs is more obvious. Therefore, the time taken for all node pairs to adjust until a relatively stable state is achieved is longer than that with smaller network topology.

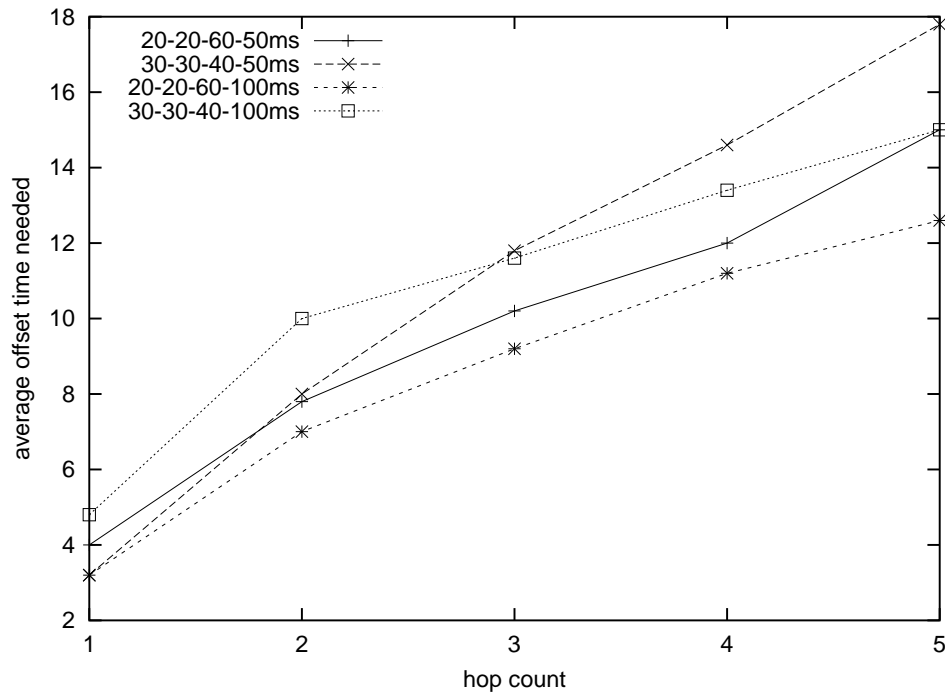


Figure 5.19: Average offset time needed for class 2 traffic with $R_{1,2}^d = 3$, and arrival rate of $0.1 \text{ bursts}/\mu\text{sec}$.

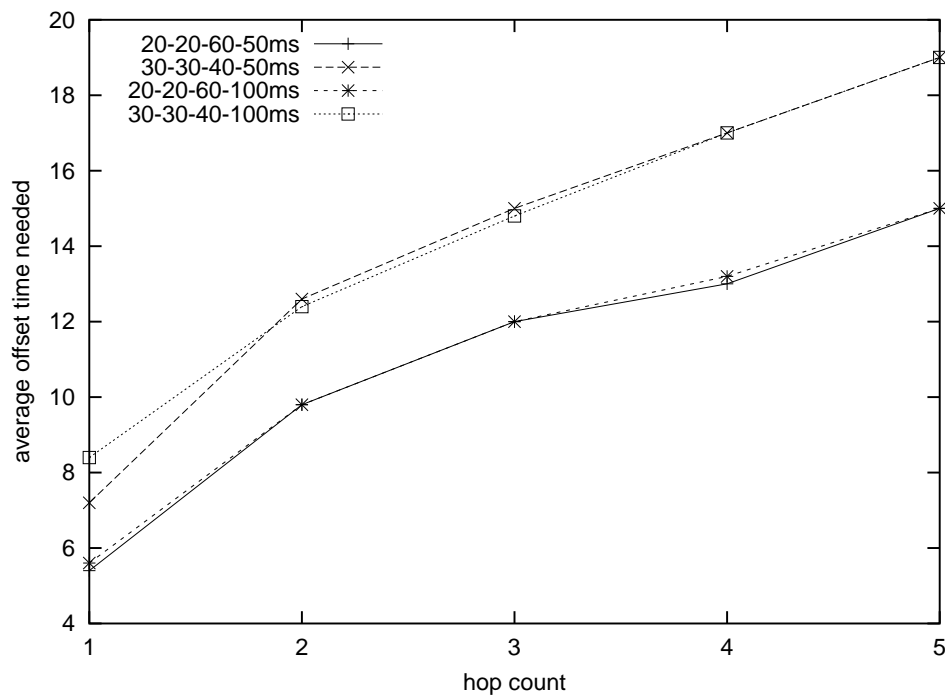


Figure 5.20: Average offset time needed for class 2 traffic with $R_{1,2}^d = 3$, and arrival rate of $0.2 \text{ bursts}/\mu\text{sec}$.

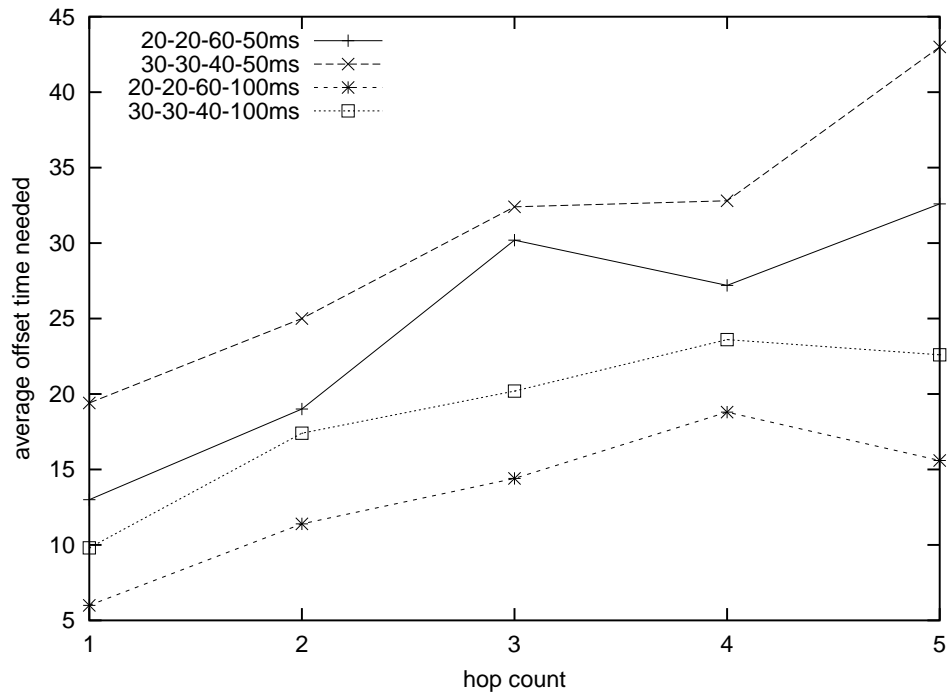


Figure 5.21: Average offset time needed for class 3 traffic with $R_{2,3}^d = 5$, and arrival rate of $0.1 \text{ bursts}/\mu\text{sec}$.

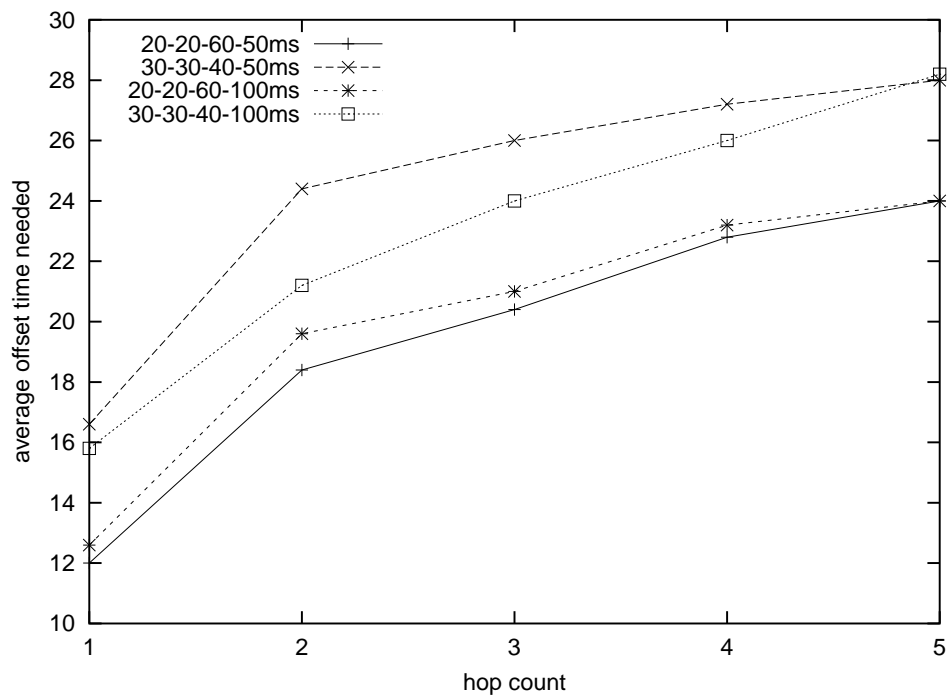


Figure 5.22: Average offset time needed for class 3 traffic with $R_{2,3}^d = 5$, and arrival rate of $0.2 \text{ bursts}/\mu\text{sec}$.

5.4 Summary

In this chapter, a feedback-based offset time selection method to provide edge-to-edge proportional QoS in optical burst switching networks has been proposed. Apart from providing edge-to-edge proportional QoS, the proposed FOTS is able to ensure fairness in terms of the proportional factor achieved for node pairs with various hop counts. FOTS is attractive since intelligent decisions are taken at the edge node rather than at the core nodes, relieving the core nodes of the processing and algorithmic burden. Also, as online link states are obtained via probe packets, FOTS inherently captures the dynamic traffic loading and topological connectivity of the network. As the probe packets are sent periodically, the signalling overhead is low. We have demonstrated the effectiveness of FOTS for different values of such parameters as the traffic measurement period, traffic proportion, traffic load, and predefined proportional ratio.

Chapter 6

Conclusions

The wavelength division multiplexing (WDM) technology which can support multiple wavelength channels, each at a rate of 10Gbps can easily offer bandwidths in the order of tens of Terabits per second. It is seemingly a promising candidate to sustain the explosive growth in the next generation IP based Optical Internet. Optical burst switching (OBS) has been considered as a viable technique to transport IP traffic over WDM considering various limitation imposed by the current optical device technology. Recently, various emerging applications such as video on demand and teleconferencing are not only bandwidth intensive, but demanding certain requirements to be met. The customers are willing to pay for more predictable service tied to service-level agreements (SLAs). It is crucial that the next generation WDM OBS network is able to provide different priorities to different types of traffic, referred to as Quality of Service (QoS), and maintain control mechanisms to enforce high performance for high priority traffic. At the same time it should be able to accommodate the legacy services without causing excessive performance degradation.

In this dissertation, we investigated various techniques on supporting QoS in the WDM OBS network such that service differentiation and fairness can be delivered. Service differentiation

is centered around the extra offset time based method as it is an attractive solution which does not require complex processing at the core nodes and hence relieving the core nodes from heavy computation burden. In the context of fairness, we stress upon two important aspects. The first one is fairness within a class where ingress egress node pairs experience similar dropping performance regardless of their path length. The second focus is on providing higher dropping performance to higher priority traffic without causing excessive dropping of the lower priority traffic.

6.1 Research Contribution

In Chapter 3, we developed new burst rescheduling algorithms as an alternative to the computationally complex void filling algorithms. Burst rescheduling techniques based on wavelength reassignment and Last-hop FDL reassignment have been proposed. Two burst rescheduling algorithms which make use of wavelength reassignment have been developed for WDM OBS networks without FDLs namely On-demand Burst Rescheduling (ODBR) and Aggressive Burst Rescheduling (ABR). Further, we developed Burst Rescheduling with Wavelength and last-hop FDL Reassignment (BR-WFR), making use of one or both of - wavelength reassignment and last-hop FDL reassignment - which is suitable for WDM OBS networks equipped with limited FDLs. It has been shown that the proposed burst rescheduling algorithms perform significantly better than existing simple LAUC algorithm in terms of burst dropping probability. At the same time their performance is close to that of the existing complex LAUC-VF algorithm at low loads. In particular, the dropping performance of the low priority traffic has been improved significantly at low loads thereby contributing to the aspect of fairness. The signaling overhead incurred by the proposed algorithms has been studied and observed to be less significant when compared to the computational complexity gain achieved over LAUC-VF.

In WDM OBS networks, bursts that traverse longer hop paths have higher chances of being dropped compared to bursts that traverse shorter hop paths resulting in fairness problem.

In Chapter 4, we developed a link scheduling state based fairness improvement method which can be used in a classless as well as a multi-class environment. The proposed link scheduling state based offset selection (LSOS) method collects link scheduling state information and uses it to determine the offset times for routes with different hop lengths. By using the online link state information, this method periodically computes and adapts the offset times needed, thus inherently accounting for the traffic loading patterns and network topological connectivity. By using the link state information on a few links only, the need for global state information can be avoided. LSOS ensures that the delay experienced by a burst is low and shorter-hop bursts are not over-penalized while improving the performance of longer-hop bursts. The performance of the proposed k -link based LSOS (k -LSOS) was studied with A-LSOS and 1-LSOS where A-LSOS uses link states on all links along a route while 1-LSOS uses only the link state on the first link. The effectiveness of the proposed method has been demonstrated for classless and multi-class environments with identical and non-identical traffic demands. Improvement in fairness is achieved with a predefined acceptable range of offset times.

In Chapter 5, we approached the problem of supporting edge-to-edge proportional QoS. Existing proportional QoS methods in WDM OBS networks are per-hop based. Supporting per-hop proportional QoS does not guarantee end-to-end proportional QoS. Per-hop approach cannot be directly extended to the multiple-hop case. We developed a feedback-based offset time selection (FOTS) method to provide edge-to-edge proportional QoS in terms of burst dropping probability. By making use of the link states collected by the probe packets sent in a periodic manner, FOTS selects appropriate offset times needed to achieve the predefined proportional QoS among different classes of traffic for various ingress-egress node pairs. Since the offset time selection is done for the node pairs, FOTS ensures fairness among node pairs

with various hop lengths in terms of achieving the predefined proportional QoS. As the on-line link state information is used, this method which periodically computes the offset times inherently captures the dynamic traffic loading and topological connectivity of the network. Further, with the probe packets sent periodically, the signalling overhead is low. An analysis on providing proportional QoS with offset time for a single link model was presented. This analysis provides some insights on how these two concepts (of proportional QoS and offset time) work together. Through the analysis, we discussed the potential of using extra offset time to provide proportional ratios in terms of the range of the achievable proportional ratio. We also studied what could affect the offset time needed to achieve a predefined proportional ratio. Finally, the effectiveness of the proposed method was evaluated through simulation experiments.

In short, we have proposed efficient QoS solutions focusing on supporting inter-class service differentiation and intra-class fairness in WDM OBS networks. First, we developed efficient burst rescheduling algorithms as an alternative to the existing complex LAUC-VF algorithm. These algorithms support service differentiation and contribute to the aspect of fairness. Next, we developed the link scheduling state based fairness method to ensure that bursts traversing different hop paths have similar performance in terms of dropping performance. Finally, we developed an edge-to-edge proportional QoS method for proportional QoS provisioning in WDM OBS networks which can also achieve fairness for node pairs with different hop counts.

6.2 Future Work

Various service differentiation schemes proposed for the WDM OBS networks aim to provide predictable, controllable and effective QoS to the users. The basic foundation for this is a stable and reliable network where failures hardly occur. In real life situation, failures are

unavoidable due to various reasons such as fiber cut or node failure. A fiber cut can cause enormous data loss due to the high traffic volume. Further, there are various emerging mission-critical applications which require guaranteed fault tolerance (i.e. the ability of the network to automatically recovers from failures) and recovery time. Therefore, one of the future extensions is to take into consideration failure while providing QoS provisioning. Substantial studies on various survivability mechanisms in WDM networks are available. However, only a little work on survivability issue for OBS networks has been developed in the literature. Specifically, developing mechanisms on providing service differentiation in the perspective of survivability still requires considerable attention.

Other than dropping probability, QoS parameters such as bandwidth and delay are anticipated to become increasingly important as the number of simultaneous bandwidth intensive connections from real-time applications increases. This leads to the issue of how to manage the resources such as bandwidth effectively so that the required QoS from different traffic classes can be delivered during different network conditions. At the same time, it is desirable to ensure high utilization of the resources. Possible extension could be on developing efficient technique on supporting QoS resource management, addressing how well the QoS scheme can provide bandwidth and latency recovery over congested networks.

Bibliography

- [1] A. Gnauck *et al.*, “One terabit/s transmission experiment,” *OSA Technical Digest Series Conference Edition, Optical Fiber Communications 96*, vol. 2, no. PD20, February 1996.
- [2] R. E. Wagner, R. C. Alferness, A. A. M. Saleh, and M. S. Goodman, “MONET: multi-wavelength optical networking,” *Journal of Lightwave Technology*, vol. 14, pp. 1349, 1996.
- [3] C. Siva Ram Murthy and G. Mohan, “WDM optical networks: concepts, design and algorithms,” *Prentice-Hall*, November 2001.
- [4] R. Ramaswami and K. N. Sivarajan, “Optical networks: a practical perspective,” second edition, *San Francisco: Morgan Kaufmann Publishers*, 2002.
- [5] T. E. Stern and K. Bala, “Multiwavelength optical networks: a layered approach,” *Reading MA: Addison-Wesley*, 1999.
- [6] K. M. Sivalingam, S. Subramaniam, “Emerging Optical Network Technologies : Architectures, Protocols and Performance,” *Springer*, 2005.
- [7] B. S. Arnaud, “Architectural and engineering issues for building an optical Internet,” *Proceedings of SPIE Conference on All-Optical Networking*, pp. 358-377, November 1998.
- [8] Sudhir Dixit, “IP over WDM: Building the next-generation optical Internet,” *Wiley*, March 2003.

-
- [9] J. Manchester, J. Anderson, B. Doshi, and S. Dravida, "IP over SONET," *IEEE Communication Magazine*, vol. 36, pp. 136-142, May 1998.
 - [10] H. M. Chaskar, S. Verma, and R. Ravikanth, "A framework to support IP over WDM using optical burst switching," *Proceedings of Optical Networks Workshop*, January 2000.
 - [11] S. Verma, H. Chaskar, and R. Ravikanth, "Optical burst switching: a viable solution for terabit IP backbone," *IEEE Network*, pp. 48-53, November/December 2000.
 - [12] C. Qiao and M. Yoo, "Optical burst switching (OBS) - a new paradigm for an optical Internet," *Journal of High Speed Networks*, vol. 8, pp. 69-84, October 1999.
 - [13] J. S. Turner, "Terabit burst switching," *Journal of High Speed Networks*, vol. 8, pp. 3-16, 1999.
 - [14] Y. Xiong, M. Vandenhoute, and H. C. Cankaya, "Control architecture in optical burst-switched WDM networks," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 10, pp. 1838-1851, October 2000.
 - [15] A. R. Moral, P. Bonenfant, and M. Krishnaswamy, "The optical Internet: architectures and protocols for the global infrastructure of tomorrow," *IEEE Communications Magazine*, vol.39, no.7, pp.152-159, July 2001.
 - [16] R. Braden *et al.*, "Integrated services in the Internet architecture: an overview," *RFC 1633*, June 1994.
 - [17] S. Blake *et al.*, "An architecture for differentiated services," *RFC 2475*, Remember 1998.
 - [18] X. Cao, J. Li, Y. Chen, and C. Qiao, "TCP/IP packets assembly over optical burst switching network," *Proceedings of GLOBECOM 2002*, vol. 3, pp. 2808-2812, September 2002.

-
- [19] J. Ramamirtham and J. Turner, "Design of wavelength converting switches for optical burst switching," *Proceedings of INFOCOM 2002*, vol. 1, pp. 362-370, 2002.
- [20] E. Varvarigos and V. Sharma, "The ready-to-go virtual circuit protocol : A loss-free protocol for multigigabit networks using FIFO buffers," *IEEE/ACM Transactions on Networking*, vol. 5, no. 5, pp. 705-718, October 1997.
- [21] I. Widjaja, "Performance analysis of burst admission-control protocols," *IEEE Proceedings of Communications*, vol. 143, no. 1, pp. 7-14, February 1995.
- [22] M. Yoo and C. Qiao, "Just-enough-time(JET): a high speed protocol for bursty traffic in optical networks," *Digest of IEEE/LEOS Summer Topical Meetings on Technologies for a Global Information Infrastructure*, pp. 26-27, August 1997.
- [23] K. Dolzer and C. Gauger, "On burst assembly in optical burst switching networks - a performance evaluation of Just-Enough-Time," *Proceedings of the 17th International Teletraffic Congress*, pp 149-161, September 2001.
- [24] C. Qiao, "Labeled Optical Burst Switching for IP-over-WDM Integration," *IEEE Communications Magazine*, pp. 104-114, September 2000.
- [25] Y. Xiong, M. Vandenhouste, and H. C. Cankaya, "Design and analysis of optical burst-switched networks," *Proceedings of SPIE 1999*, vol. 3843, pp. 12-119, September 1999.
- [26] M. Yoo and C. Qiao, "A new optical burst switching (OBS) protocol for supporting quality of services, " *Proceedings of SPIE 1998*, vol. 3531, pp. 396-405, November 1998.
- [27] C. Qiao and M. Yoo, "Choices, features and issues in optical burst switching (OBS)," *Optical Network Magazine*, vol. 1, no. 2, pp. 36-44, April 2000.
- [28] M. Yoo and C. Qiao, "QoS performance in IP over WDM networks," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 10, pp. 2062-2071, October 2000.

-
- [29] V. Vokkarane and J. Jue, "Prioritized burst segmentation and composite burst assembly techniques for QoS support in optical burst-switched networks," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 7, pp. 1198-1209, September 2003.
- [30] M. Yang, S. Q. Zheng, D. Verchere, "A QoS supporting scheduling algorithm for optical burst switching DWDM networks," *Proceedings of IEEE GLOBECOM 2001*, vol.1, pp. 86-91, 2001.
- [31] A. Kaheel and H. Alnuweiri, "Strict priority scheme for quality-of-service provisioning in optical burst switching networks," *Proceedings of ISCC 2003*, vol. 1, pp. 16-21, June 2003.
- [32] L. Yang, Y. Jiang, and S. Jiang, "A probabilistic preemptive scheme for providing service differentiation in OBS networks," *Proceedings of IEEE GLOBECOM 2003*, vol. 5 pp. 2689-2693, December 2003.
- [33] Y. Chen, M. Hamdi, and D. H. K. Tsang, "Proportional QoS over OBS networks," *Proceedings of IEEE GLOBECOM 2001*, vol. 3, pp. 1510-1514, 2001.
- [34] Y. Chen, C. Qiao, M. Hamdi, and D. H. K. Tsang, "Proportional differentiation: A scalable QoS approach," *IEEE Communications Magazine*, vol. 41, no. 6, pp. 52-58, June 2003.
- [35] H. C. Cankaya, S. Charcranoon, and T. S. El-Bawab, "A preemptive scheduling technique for OBS networks with service differentiation," *Proceedings of IEEE GLOBECOM 2003*, vol. 5, pp. 2704-2708, December 2003.
- [36] W. Liao and C. Loi, "Providing service differentiation for optical-burst-switched networks," *Journal of Lightwave Technology*, vol. 22, no. 7, pp. 1651-1660, July 2004.
- [37] C. Dovrolis, D. Stiliadis, and P. Ramanathan, "Proportional differentiated services: delay differentiation and packet scheduling," *IEEE/ACM Transactions on Networking*, vol. 10, no. 1, pp. 12-26, February 2002.

-
- [38] C. Dovrolis and P. Ramanathan, "Proportional differentiated services, part II: loss rate differentiation and packet dropping," *Proceedings of IWQOS 2000*, pp. 53-61 5-7, June 2000.
- [39] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," *IEEE/ACM Transactions on Networking*, vol. 1, no. 4, pp. 397-413, August 1993.
- [40] Q. Zhang, V. Vokkarane, Biao Chen, and Jason Jue, "Early drop and wavelength grouping schemes for providing absolute QoS differentiation in optical burst-switched networks," *Proceedings of IEEE GLOBECOM 2003*, vol. 5, pp. 2628-2632, December 2003.
- [41] M. H. Phung, K. C. Chua, G. Mohan, M. Motani, and T. C. Wong, "A preemptive differentiation scheme for absolute loss guarantees in OBS networks," *Proceedings of IASTED OCSN 2004*, July 2004.
- [42] A. Kumar, J. Kaur, and H. M. Vin, "End-to-End proportional loss differentiation," *Technical report TR-01-33*, University of Texas Austin, September 2001.
- [43] H. Liu, "IP over WDM," *Wiley*, September 2002.
- [44] V. Vokkarane, Q. Zhang, J. Jue, and B. Chen, "Generalized burst assembly and scheduling techniques for QoS support in optical burst-switched networks," *Proceedings of GLOBECOM 2002*, vol. 3, pp. 2747-2752, September 2002.
- [45] J. Luo, Q. Zeng, H. Chi, Z. Zhang, and H. Zhao "The impacts of burst assembly on the traffic properties in optical burst switching networks," *Proceedings of ICCT 2003*, vol. 1, pp.521-524, April 2003.
- [46] R. Rajaduray, D.J. Blumenthal, and S. Ovadia, "Impact of burst assembly parameters on edge router latency in an optical burst switching network," *Proceedings of IEEE LEOS 2003*, vol. 1, pp. 55-56, October 2003.

-
- [47] J. Liu and N. Ansari, "The impact of the burst assembly interval on the OBS ingress traffic characteristics and system performance," *Proceedings of ICC 2004*, vol. 3, pp. 1559-1563, June 2004.
- [48] J. Y. Wei and R. I. McFarland, "Just-in-time signaling for WDM optical burst switching networks," *Journal of Lightwave Technology*, vol. 18, no. 12, pp. 2019-2037, December 2000.
- [49] J. Y. Wei, J. L. Pastor, R. S. Ramamurthy, and Y. Tsai, "Just-in-time optical burst switching for multiwavelength networks," *Proceedings of IFIP Broadband Communications*, pp. 339-352, November 1999.
- [50] E. A. Varvarigos and V. Sharma, "Ready-to-go virtual circuit protocol: A loss-free protocol for multi- gigabit networks using FIFO buffers," *IEEE/ACM Transactions on Networking*, vol. 5, no. 5, pp. 705-718, October 1997.
- [51] P. Kermani and L. Kleinrock, "Virtual cut-through: a new computer communication switching technique," *Computer Networks Magazine*, vol. 3 no. 3 pp.267-286, September 1979.
- [52] I. Baldine, M. Cassada, A. Bragg, G. Karmous-Edwards, and D. Stevenson, "Just-in-time optical burst switching implementation in the ATDnet all-optical networking testbed," *Proceedings of IEEE GLOBECOM 2003*, vol. 5, pp. 2777-2781, December 2003.
- [53] J. Teng and G. N. Rouskas, "A comparison of the JIT, JET, and Horizon wavelength reservation schemes on a single OBS node," *Proceedings of the First International Workshop on Optical Burst Switching*, October 2003.
- [54] A. Detti, V. Eramo, and M. Listanti, "Performance evaluation of a new technique for IP support in a WDM optical network: optical composite burst switching (OCBS)," *Journal of Lightwave Technology*, vol. 20 no. 2 pp.154-165, February 2002.

-
- [55] V. Vokkarane, G. P. V. Thodime, V. U. B. Challagulla, and J. Jue, "Channel scheduling algorithms using burst segmentation and FDLs for optical burst-switched networks," *Proceedings of ICC 2003*, vol. 2, pp. 1443-1447, May 2003.
- [56] M. Neuts, Z. Rosberg, H. L. Vu, J. White, and M. Zukerman, "Performance enhancement of optical burst switching using burst segmentation," *Proceedings of ICC 2003*, vol. 3, pp. 1828-1832, May 2003.
- [57] Y. Chen, H. Wu, D. Xu, and C. Qiao, "Performance analysis of optical burst switched node with deflection routing," *Proceedings of ICC 2003*, vol. 2, pp. 1355-1359, May 2003.
- [58] F. Farahmand and J. Jue, "Look-ahead window contention resolution in optical burst switched networks," *Proceedings of IEEE HPSR 2003*, pp. 147 - 151, June 2003.
- [59] S. Y. Wang, "Using TCP congestion control to improve the performances of optical burst switched networks," *Proceedings of ICC 2003*, vol. 2, pp. 1438-1442, May 2003.
- [60] S. Gowda, R. K. Shenai, K. M. Sivalingam, and H.C. Cankaya, "Performance evaluation of TCP over optical burst-switched (OBS) WDM networks," *Proceedings of ICC 2003*, vol. 2, pp. 1433-1437, May 2003.
- [61] X. Wang, A. Saito, H. Morikawa, and T. Aoyama, "Priority-based wavelength assignment algorithm for burst switched photonic networks," *Proceedings of OFC 2002*, pp.765-767, March 2002.
- [62] J. Li and C. Qiao, "Schedule burst proactively for optical burst switching networks," *Proceedings of IEEE GLOBECOM 2003*, vol. 5, pp.2787-2791, December 2003.
- [63] J. Xu, C. Qiao, J. Li, and G. Xu, "Efficient channel scheduling algorithms in optical burst switched networks," *Proceedings of INFOCOMM 2003*, vol. 3, pp. 2268-2278, 2003.

-
- [64] M. Iizuka, M. Sakuta, Y. Nishino, and I. Sasase, "A scheduling algorithm minimizing voids generated by arriving bursts in optical burst switched WDM network," *Proceedings of IEEE GLOBECOM 2002*, vol. 3, pp. 2736-2740, November 2002.
- [65] H. Zhang, "Service disciplines for guaranteed performance service in packet-switching networks," *Proceedings of the IEEE*, vol. 83, no. 10, pp. 1374-1396, October 1995 .
- [66] M. Yoo and C. Qiao, "Supporting Multiple Classes of Services in IP over WDM Networks," *Proceedings of IEEE GLOBECOM 1999*, pp. 1023-1027, December 1999.
- [67] N. Barakat and E. H. Sargent, "The influence of low-class traffic load on high-class performance and isolation in optical burst switching systems," *Proceedings of IEEE ICC 2004*, vol. 3, pp. 1554-1558, June 2004.
- [68] K. Dolzer, C. Gauger, J. Späth, and S. Bodamer, "Evaluation of reservation mechanisms for optical burst switching," *AEÜ International Journal of Electronics and Communications*, vol. 55, no. 1, pp. 1826, 2001.
- [69] W. So, Y. Cha, S. Roh, and Y. Kim, "Offset time decision (OTD) algorithm for guaranteeing the requested QoS of high priority traffic in OBS networks," *Proceeding of APOC 2001*, pp. 286-296, November 2001.
- [70] S. Kim, A. Choi, and M. Kang, "Performance analysis for prioritized multi-classes in optical burst switching networks," *Proceedings of ICACT 2004*, vol. 1, pp. 72-74, February 2004.
- [71] N. Barakat and E. H. Sargent, "An accurate model for evaluating blocking probabilities in multi-class OBS systems," *IEEE Communications Letters*, vol. 8, no. 2, pp. 119-121, February 2004.
- [72] M. Neuts, H. Vu, and M. Zukerman, "Performance analysis of optical composite burst switching," *IEEE Communications Letters*, vol. 6, no. 8 pp.346-348, August 2002.

-
- [73] Z. Rosberg, H.L. Vu, and M. Zukerman, "Burst segmentation benefit in optical switching," *IEEE Communications Letters*, vol. 7, no. 3, pp 127-129, March 2003.
- [74] K. Long, R. S. Tucker, C. Wang, "A new framework and burst assembly for IP DiffServ over optical burst switching networks," *Proceedings of GLOBECOM 2003*, vol. 6, pp. 3159 - 3164, December 2003.
- [75] F. Poppe, K. Laevens, and H. Michiel, "Quality-of-service differentiation and fairness in optical burst-switched networks," *Proceedings of OPTICOMM 2002*, vol. 4874, pp. 118-124, 2002.
- [76] A. Elwalid, C. Jin, S. Low, and I. Widjaja, "MATE: MPLS adaptive traffic engineering," *Proceedings of INFOCOM 2001*, vol.3, pp. 1300-1309, April 2001.
- [77] M. Elhaddad, R. Melhem, T. Znati, and D. Basak, "Traffic shaping and scheduling for OBS-based IP/WDM Backbones," *Proceedings of OPTICOMM 2003*, pp. 357-368, October 2003.
- [78] J. Li, G. Mohan, and K. C. Chua, "Load balancing using adaptive alternate routing in IP-over-WDM optical burst switching networks," *Proceedings of OPTICOMM 2003*, pp 336-345, October 2003.
- [79] K. Thompson, G. Miller, and R. Wilder, "Wide area Internet traffic patterns and characteristics," *IEEE Network Magazine*, vol. 11, pp.10-23, November-December 1997.
- [80] C. Fraleigh *et al.*, "Packet-level traffic measurements from the sprint IP backbone," *IEEE Network Magazine*, vol. 17, pp. 6-16, November-December 2003.
- [81] V. Paxson and S. Floyd, "Wide-area traffic: The failure of Poisson modeling," *Proceedings of ACM SIGCOMM 1994*, pp. 257-268, August 1994.

-
- [82] Q. Zhang, V. Vokkarane, B. Chen, and J. Jue, "Early drop scheme for providing a absolute QoS differentiation in optical burst-switched networks," *Proceedings of IEEE HPSR 2003*, pp. 153-157, June 2003.
- [83] L. Xu, H. G. Perros, and G. Rouskas, "Techniques for optical packet switching and optical burst switching," *IEEE Communications Magazine*, pp. 136-142, January 2001.
- [84] C. Guillemot *et al.*, "Transparent optical packet Switching: The European ACTS KEOPS project approach," *IEEE/OSA Journal of Lightwave Technology*, vol. 16, no. 12, pp. 2117-2134, December 1998.
- [85] D. K. Hunter *et al.*, "WASPNET: A wavelength switched packet network," *IEEE Communications magazine*, vol. 37, no. 3, pp. 120-129, March 1999.
- [86] F. Callegati, H. Cankaya, Y. Xiong, and M. Vandenhoute, "Design issues of optical IP routers for internet backbone applications, " *IEEE Communications magazine*, pp. 124-128, December 1999.
- [87] R. Jain and D. Chiu, "Analysis of the increase/decrease algorithms for congestion avoidance in computer networks," *Journal of Computer Networks*, no.17, pp.1-14, 1989.

Author's Publications

- [1] S. K. Tan, G. Mohan, and K. C. Chua, "Algorithms for Burst Rescheduling in WDM Optical Burst Switching Networks," *Computer Networks - International Journal of Computer and Telecommunications Networking*, vol. 41, no. 1, pp. 41-55, January 2003.
- [2] S. K. Tan, G. Mohan, and K. C. Chua, "Burst Rescheduling with Wavelength and Last-hop FDL Reassignment in WDM Optical Burst Switching Networks," *Proceedings of IEEE International Conference on Communications (ICC) 2003*, vol. 26, no. 1, pp. 1448 - 1452, May 2003.
- [3] S. K. Tan, G. Mohan, and K. C. Chua, "Link Scheduling State Information Based Offset Management for Fairness Improvement in WDM Optical Burst Switching Networks," *Computer Networks - International Journal of Computer and Telecommunications Networking*, vol. 45, no. 6, pp. 819-834, August 2004.
- [4] S. K. Tan, G. Mohan, and K. C. Chua, "Feedback-based Offset Time Selection Method on Supporting End-to-End Proportional QoS in WDM Optical Burst Switching Networks," (to be submitted).